# Lecture 1: Overview

## 1 Diffusion by discontinous movements

In 1827 Robert Brown, observed that suspended pollen grains are in an uninterrupted and irregular "swarming" motion. Brown was a botanist and at first he believed that only organic materials exhibited this agitation. But very soon he extended his observations to particles of inorganic material, such as a ground-up fragment of the Sphinx. Through the nineteenth century there was a intermittent discussion concerning the cause of this *Brownian motion*, and in 1877 Delsaux suggested that the impact of molecules on a macroscopic particle produces observable displacements. In 1905, after nearly a century of debate, Einstein definitively explained this phenomenon [6, 7].

### 1.1 Einstein's derivation of the diffusion equation

Our interest here is in Einstein's derivation of the diffusion equation, which is very different from that of Fourier. We consider one-dimensional Brownian motion by projecting the location of the particle onto a straight line which we call the $x$-axis.

Einstein's assumptions are the following: (i) the particles move independently of one another; (ii) we observe particle positions at time intervals $\tau$ which are much greater than the time intervals between molecular collisions. As a result, the motion in one interval is independent of what happened in the previous interval.

In the interval $\tau$ each particle has a random displacement $\Delta$ along the $x$-axis. The probability density function (PDF) of $\Delta$ is $\phi(\Delta)$. This means that if we observe $N \gg 1$ particles for a time $\tau$ then the number of particles which are displaced through a distance which lies between $\Delta$ and $\Delta + \mathrm{d}\Delta$ is

$$\mathrm{d}N = N\phi(\Delta)\,\mathrm{d}\Delta. \tag{1}$$

The PDF $\phi(\Delta)$ does not change from interval to interval, and $\phi$ is symmetric and normalised:

$$\phi(\Delta) = \phi(-\Delta)\,, \qquad \int_{-\infty}^{\infty} \phi(\Delta)\,\mathrm{d}\Delta = 1. \tag{2}$$

The symmetry of $\phi$ implies that the displacements are unbiased. The average of any function of $\Delta$, $f(\Delta)$, is

$$\bar{f} \equiv \int_{-\infty}^{\infty} f(\Delta)\phi(\Delta)\,\mathrm{d}\Delta\,. \tag{3}$$

In particular, $\overline{\Delta^2}$ is the mean square displacement in a single step.

If the concentration of particles at time $t$ is denoted by $c(x,t)$, then the evolution of $c$ is determined from the *master equation*:

$$c(x, t+\tau) = \int_{-\infty}^{\infty} c(x-\Delta, t)\phi(\Delta)\,\mathrm{d}\Delta\,. \tag{4}$$


Figure 1: Simulated Brownian motion using MATLAB; the routine rand is used to generate a sequence of 200 random displacements.

The integral over $\Delta$ is a sum over the prior locations at time $t$ of the particles that are at $x$ at time $t+\tau$. Thus, the number of particles in the interval $(x-\Delta, x-\Delta+\mathrm{d}\Delta)$ is $c(x-\Delta,t)\mathrm{d}\Delta$ and $\phi(\Delta)$ is the fraction of these particles which jump from $x-\Delta$ onto $x$.

If the concentration $c(x,t)$ changes on a length scale which is much greater than the root mean square displacement, then we can approximate the integral equation (4) by the diffusion equation. This assumption that $c$ is slowly varying means that it it is sensible to use a Taylor series expansion

$$c(x,t)+\tau c_t(x,t) \approx \int_{-\infty}^{\infty} \phi(\Delta)\left[c(x,t)-\Delta c_x(x,t)+\frac{\Delta^2}{2}c_{xx}(x,t)\right]\mathrm{d}\Delta. \tag{5}$$

Next, using (2), we reduce (5) to

$$c_t(x,t) \approx D c_{xx}(x,t)\,, \qquad D \equiv \frac{\overline{\Delta^2}}{2\tau}. \tag{6}$$

This is the diffusion equation, and $D$ is the diffusivity.

The greatness of Einstein's contribution to this subject is not the derivation above but rather his formula for the diffusivity of a macroscopic particle

$$D = \frac{RT}{6\pi N_a \nu a}, \tag{7}$$

where $R$ is the gas constant, $T$ the absolute temperature, $N_a$ the Avogadro number, $\nu$ the coefficient of viscosity and $a$ the radius of the particle. Coincidentally, (7) was also discovered in 1905 by William Sutherland in Australia. This relation enabled Perrin to determine Avogadro's number by observing Brownian displacements [7].

The diffusion equation is an approximation of the more exact master equation. As we try to design parameterizations of nonlocal mixing processes, in which scale separation assumptions are shaky, we should pay more attention to this history and consider the possibility of using integral equations such as (4). Notice also that if the Taylor expansion in (5) is continued to higher order then one will usually (i.e. for most kernels $\phi$) obtain a hyperdiffusive term such as $c_{xxxx}$.

## 1.2 The method of moments

As a check on the derivation of (6), we take a different approach using the method of moments. A *moment* of the concentration is an integral of the form

$$\int_{-\infty}^{\infty} x^n c(x,t)\,\mathrm{d}x. \tag{8}$$

The zeroth moment, $n = 0$ in (8), is the total number of particles:

$$N = \int_{-\infty}^{\infty} c(x,t)\,\mathrm{d}x. \tag{9}$$

The first and second moments can be interpreted as the center of mass and moment of inertia of the concentration profile.

We expect that $N$ is constant, and it is educational to verify this conservation law for both the master equation and the diffusion equation by "taking the zeroth moment". Integrating (4) from $x = -\infty$ to $x = +\infty$, and changing the order of the integrals on the right-hand side gives

$$N(t+\tau) = \int_{-\infty}^{\infty} \mathrm{d}\Delta\, \phi(\Delta) \int_{-\infty}^{\infty} \mathrm{d}x\, c(x-\Delta,t). \tag{10}$$

Changing variables to $x' = x - \Delta$ in the inner integral, and using (2), gives the particle conservation law $N(t+\tau) = N(t)$. The diffusive analog of particle conservation is easily obtained by integrating the diffusion equation (6) from $x = -\infty$ to $x = +\infty$. Provided that $Dc_x$ vanishes at $x = \pm\infty$ (physically, there is no flux of particles from infinity), one immediately finds that $N_t = 0$.

Extending the procedure above to higher moments, we can make a comparison between the exact results for $\int x^n c\,\mathrm{d}x$ and the diffusive approximation of these same integrals. To take the first moment of the diffusion equation, multiply (6) by $x$ and integrate from $x = -\infty$ to $x = +\infty$. Once again, we use integration by parts and assume that terms such as $xc_x$ and $c$ vanish as $x \to \pm\infty$. Thus we find that the center of mass is stationary

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{-\infty}^{\infty} xc(x,t)\,\mathrm{d}x = 0. \tag{11}$$

The same result can be obtained by taking the first moment of the master equation. The center of mass is stationary because in (2) we assume that the PDF of displacements is symmetric.

Continuing, we come to the second moment. For the diffusion equation we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} x^2 c\,\mathrm{d}x = 2D\int_{-\infty}^{\infty} c\,\mathrm{d}x = 2DN, \tag{12}$$

where, as before, the terms which fall outside the integration by parts are zero because of the rapid decay of $c$ as $x \to \pm\infty$. The student should show that from the master equation

$$\int_{-\infty}^{\infty} x^2 c(x,t+\tau)\,\mathrm{d}x - \int_{-\infty}^{\infty} x^2 c(x,t)\,\mathrm{d}x = \int_{-\infty}^{\infty} \Delta^2 \phi(\Delta)\,\mathrm{d}\Delta. \tag{13}$$

Recalling the definition of the diffusivity in (6), we see that in the limit $\tau \to 0$ the difference equation in (13) can be approximated by the differential equation in (12).

The law in (12), that the mean square displacement of a cloud of particles grows linearly with time, is often taken to be the defining characteristic of diffusion. As we will see later, there are dispersive processes which have other power-laws, such as $\int x^2 c\,\mathrm{d}x \propto t^{1/2}$. These processes are referred to as "anomalous diffusion".

# 2 Diffusion by continuous movements

## 2.1 Lagrangian time series

In 1922 Taylor [11] analyzed the diffusing power of a velocity field. The basic concept here is that of a Lagrangian time series, such as the $x$-velocity of a tagged fluid particle, $u(t)$, as a function of time. This data is Lagrangian (i.e., following a "float"), *not* Eulerian (i.e, obtained from a "current meter" fixed in space). The velocity time series might look like figure 2. Clearly there is some regularity: evenly spaced maxima and minima are obvious, and we might guess that there is a wave which is producing oscillatory displacements. At the same time, the velocity is not completely predictable, and there is no obvious law by which we can anticipate all details of the future using observations of the past.

The simplest assumption we can make to analyze the process in figure 2 is that the velocity is statistically *stationary*. This means that average properties of the velocity, such as the mean square velocity, are not changing with time. In operational terms, the assumption of stationarity means that if we take nonoverlapping and well-separated subsamples of the time series in figure 2 then the statistical properties of the subsamples are identical.

If the time series is long enough we can chop it into $N$ chunks, each of length $T$. We define an ensemble average by considering each of the $N$ chunks as a single realization of a random process. This procedure introduces the additional assumptions that there is a decorrelation time $\tau \ll T$, and that time averages are equivalent to ensemble averages. Thinking of dispersion, Taylor imagined that each chunk was an independent particle, labeled $n = 1, 2, ..., N$, executing continuous movements. "Continuous" in this context means that the velocity of particle $n$, $u_n(t)$, is a relatively smooth function of time, at least in comparison with the jittery motion in figure 1.

We denote the position of particle $n$ by $x_n(t)$, so that if all the particles begin at $x = 0$ then

$$\frac{\mathrm{d}x_n}{\mathrm{d}t} = u_n(t), \qquad \Longrightarrow \qquad x_n(t) = \int_0^t u_n(t')\,\mathrm{d}t'. \tag{14}$$

We use angular brackets $\langle\rangle$ to denote the ensemble average. As an example of this notation, the average velocity of the $N$ particles is

$$\langle u \rangle \equiv \frac{1}{N}\sum_{n=1}^{N} u_n(t). \tag{15}$$

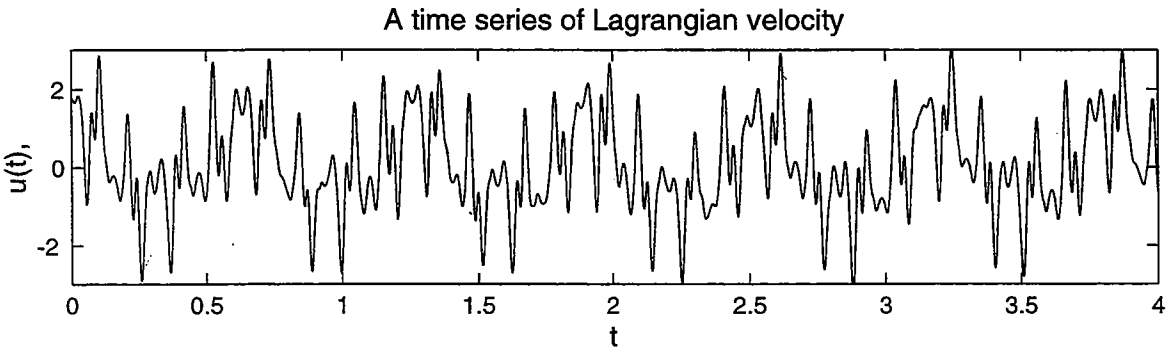


Figure 2: A time series with a spectral peak.

Because of the stationarity assumption, $\langle u \rangle$ is independent of time, and we can refer all displacements relative to the position of the center of mass by writing $x' = x - \langle u \rangle t$ and $u'_n = u_n - \langle u \rangle$. To save decorating all our subsequent $x$'s and $u$'s with primes we now assume that $\langle u \rangle = 0$.

## 2.2 Taylor's formula

The simplest measure of dispersion about the center of mass is the mean square displacement, $\langle x^2 \rangle$. We can calculate the rate of change of this quantity by first noting that:

$$\frac{\mathrm{d}x_n^2}{\mathrm{d}t} = 2x_n u_n, \qquad \text{and (14)} \Longrightarrow \qquad \frac{\mathrm{d}x_n^2}{\mathrm{d}t} = 2\int_0^t u_n(t)u_n(t')\,\mathrm{d}t'. \tag{16}$$

We now ensemble average (16). Because of stationarity, $\langle u(t)u(t') \rangle$ depends only on the time difference $t - t'$. Thus, we introduce the correlation function

$$\mathcal{C}(t - t') \equiv \langle u(t)u(t') \rangle, \tag{17}$$

and, after a change of variables, write the ensemble average of (16) as

$$\frac{\mathrm{d}\langle x^2 \rangle}{\mathrm{d}t} = 2\int_0^t \mathcal{C}(t')\,\mathrm{d}t'. \tag{18}$$

Equation (18) is Taylor's formula, which relates the variance in particle displacement $\langle x^2 \rangle$ to an integral of the Lagrangian velocity autocorrelation function $\mathcal{C}(t)$.

In the simplest situations the correlation function $\mathcal{C}(t)$ decreases rapidly to zero as $t \to \infty$ so that the integral in (18) converges. In this case, the dispersion of the ensemble at large times is characterized by a diffusivity $\langle x^2 \rangle \sim 2Dt$, where the diffusivity $D$ is related to the correlation function by:

$$D = \int_0^\infty \mathcal{C}(t)\,\mathrm{d}t. \tag{19}$$

In statistical physics, (19) is known as the Green-Kubo formula.

Taylor did not claim that turbulent dispersion was governed by the diffusion equation, (6). We will return to this point later. For the moment notice that (6) is an approximation valid only for sufficently long times that the integral in (18) has converged to the constant $D$. This restriction is related to Einstein's assumption that particle positions are observed at time intervals $\tau$ which are much greater than the decorrelation time.

# 3 Diffusion and anomalous diffusion

In the previous sections we emphasized that the diffusion equation (6) is only valid on times long compared to the decorrelation time $\tau$, and only if the concentration $c(x,t)$ varies on length scales greater than the width of the density $\phi(\Delta)$. These assumptions of scale separation in both time and space are often not satisfied in real flows. Thus, dispersion experiments over the last ten years have revealed behaviours which are much richer than those suggested by the arguments of Einstein and Taylor. Experiments often show that the growth of variance is described by a power law

$$\langle x^2 \rangle \propto t^{\xi}. \tag{20}$$

In some cases $\xi = 1$ (diffusion), but sometimes $\xi \neq 1$, in which case the process is referred to as *anomalous diffusion.*

## 3.1 Rayleigh-Bénard convection

As an example of hydrodynamic diffusion ($\xi = 1$) and transient subdiffusion ($\xi = 2/3$) we mention the experiments of Solomon and Gollub [9,8] on the dispersion of passive scalar (either methylene blue or uranine dye, or small latex spheres) along a chain of Rayleigh-Bénard convection cells (see figure 3). We refer to the passive scalar generically as "tracer".

Following the experimental procedure in figure 3, suppose that all of the tracer is initially released in a single cell. The main question is: how many cells, $N(t)$, have been invaded by tracer at time $t$? If this dispersive process is described by diffusion then we expect that $N(t) \propto t^{1/2}$. With certain interesting restrictions, this $t^{1/2}$-law is the experimental result.

The Rayleigh-Bénard flow can be approximately described using a two-dimensional and incompressible velocity field, $(u,v)$, obtained from the streamfunction

$$\psi = k^{-1} A \sin\left[k\left(x + B\sin\omega t\right)\right] W(z)\,, \qquad (u,v) = (-\psi_y, \psi_x)\,. \tag{21}$$

The parameter $A$ controls the amplitude of the flow, $k = 2\pi/\lambda$ is the wavenumber, and $W(z)$ is a function which satisfies the no-slip boundary conditions at $z = 0$ and $z = H$. The term $B\sin\omega t$ is a simple model of the lateral oscillation of the roll pattern which results from an instability which occurs when the convection is driven sufficiently strongly. Because the flow in (21) is simple, highly structured and deterministic, this is not an example of turbulent dispersion. Nonetheless, the experimental results can be summarized using the notion of an *effective diffusivity.*

The Péclet number is

$$P \equiv \frac{A}{k\kappa}, \tag{22}$$

where $\kappa$ is the molecular diffusivity of the tracer, is a nondimensional parameter which measures the importance of molecular diffusivity to advection. The Péclet number can be considered as the ratio of the time it takes a molecule to orbit around a convection cell to the diffusion time across a cell. In the experiments described here, $P$ is large and molecules make many circuits around a convection cell before Brownian motion jostles them through a distance as large as $k^{-1}$.

There are two cases which must be carefully distinguished:

***Steady rolls*** The rolls are steady if either $\omega = 0$ or $B = 0$ in (21). In either case, tracer can pass from one roll to a neighbour only via molecular diffusion. But, because molecules are advected through a distance $k^{-1}$, the dye is transported along the array of cells with an effective diffusivity $D_{\text{eff}} \propto \sqrt{A\kappa/k} \gg \kappa$. Because $D_{\text{eff}} \to 0$ if $\kappa \to 0$, the transport is limited by molecular diffusion.

***Unsteady rolls*** If $B$ and $\omega$ are both nonzero then advection (rather than molecular diffusion) can take particles through the time-averaged position of the cell boundaries. In this case, there is the possibility of transport unlimited by weak molecular diffusion.

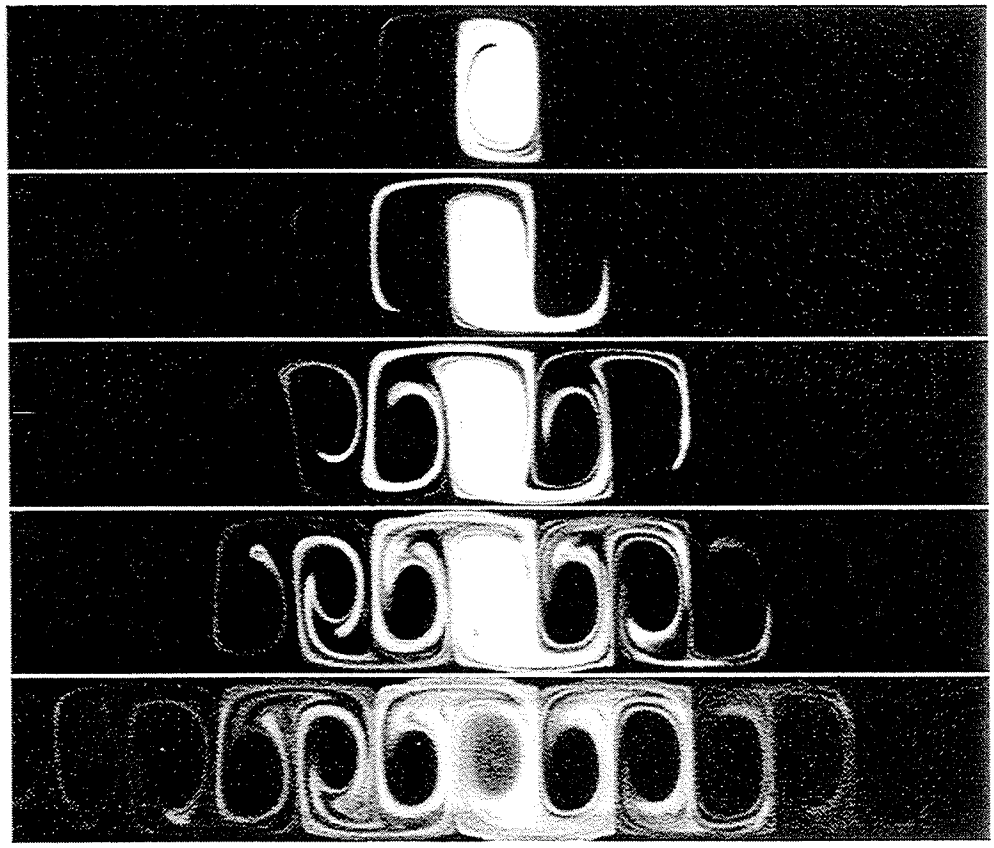

Figure 3: Transport of uranine dye along an array of convection cells with $kB = 0.12$; time (from the top): 1, 2, 4 and 10 periods of oscillation. (Figure courtesy of Tom Solomon [10].)

In the unsteady case, Solomon and Gollub show that trajectories of particles computed with the model streamfunction (21) are similar to the patterns observed experimentally. In both the numerics and the experiments, provided that $\omega B \neq 0$, the transport of particles along the array of cells (in the $x$–direction) is due to chaotic advection in the neighbourhood of the roll boundaries. This process is strikingly shown in figure 3.

A rough summary of the results is that in both the steady and the unsteady cases the dye spreads via a one-dimensional diffusive process, $\xi = 1$ in (20), with a local effective diffusivity $D_{\text{eff}}$. The number of invaded cells is $N(t) \propto \sqrt{D_{\text{eff}} t}$. In the unsteady case $D_{\text{eff}}$ is independent of the molecular diffusivity $\kappa$, while in the steady case $D_{\text{eff}} \propto \sqrt{\kappa}$. The effective diffusivity in the unsteady case is enhanced by 1 to 3 orders of magnitude over the effective diffusivity of the steady case (which in turn is much greater than the molecular diffusivity, $\kappa$).

The summary in the previous paragraph omits many interesting details. One of the more important caveats is that the effective diffusivity in the *s*teady case only describes the dispersion process at very long times:

$$N(t) \propto t^{1/2} \quad \text{when} \quad t \gg \frac{1}{k^2 \kappa}. \tag{23}$$

The time $1/k^2\kappa$ is an estimate of the time taken for molecular diffusion to transport tracer through a distance of order $k^{-1}$, from the edge of a cell to the center[1]. In this long time limit, the evolution

[1] We assume that the aspect ratio of the cells is of order unity, $kH = O(1)$.

of the tracer is slower than the intracellular diffusion time $1/k^2\kappa$ and consequently the concentration is uniform within each roll. The concentration changes rapidly at diffusive boundary layers (with thickness proportional to $\kappa^{1/2}$) which are located at the roll boundaries. The intercellular flux across these boundary layers is responsible for the spread of the tracer from one roll to the next.

The scenario described above does not have time to become established until $t \gg 1/k^2\kappa$. When $t \ll 1/k^2\kappa$ there is still a significant dispersion of tracer through many cells which is described by the anomalous diffusion law

$$N(t) \propto t^{1/3} \quad \text{when} \quad t \ll \frac{1}{k^2\kappa}\,. \tag{24}$$

The anomalous process above relies on molecular diffusion passing tracer quickly across the cell boundaries before there has been time to reach the center of newly invaded cells [3,4,13]. Thus there is a transient regime of subdiffusion which preceeds the final asymptotic diffusive law in (23).

## 3.2 Anomalous diffusion in two-dimensional turbulence

Cardoso *et al.* [2] conducted an experimental study of dispersion in a quasi-two-dimensional turbulent flow. The experimental apparatus is a shallow pan of fluid, 30cm by 30cm, and 3mm deep. The pan is filled with salty water and flow is driven electromagnetically ($\boldsymbol{E} \times \boldsymbol{B}$ forcing). The forcing is arranged so that the basic flow is a square lattice of $30 \times 30$ counter-rotating vortices. This flow is almost two-dimensional because of the large disparity between the horizontal dimensions (30 cm) and the vertical dimension (3 mm).

Although the forcing produces a regular array of vortices, this simple pattern is unstable and a two-dimensional turbulent flow emerges. Visualization of the turbulence, using tracer particles, shows that in the statistically equilibrated state there is a population of vortices whose size is two or three times the injection scale of the forcing. Each vortex emerges, moves, merges with other vortices, and eventually disappears.

Cardoso *et al.* [2] injected dye into this vortex mess and observed the two-dimensional dispersion of the dye in the horizontal plane. To measure the growth of the dye blob, they defined

$$R_m \equiv \int \sqrt{x^2+y^2}\, c(x,y,t)\,\mathrm{d}x\,\mathrm{d}y \Big/ \int c(x,y,t)\,\mathrm{d}x\mathrm{d}y\,, \tag{25}$$

and

$$R_g \equiv \sqrt{\int (x^2+y^2)\, c(x,y,t)\,\mathrm{d}x\,\mathrm{d}y \Big/ \int c(x,y,t)\,\mathrm{d}x\mathrm{d}y}\,. \tag{26}$$

The experimental scaling law is

$$(R_g, R_m) \sim t^{0.32\pm0.04}\,. \tag{27}$$

The exponent $0.32 \neq 1/2$ indicates anomalous diffusion — specifically subdiffusion, because the dispersion is slower than diffusion.

By examining typical particle trajectories, such as the one in figure 4, Cardoso *et al.*explained the subdiffusive growth in terms of an *interrupted random walk*. Consider a random walker who pauses between steps. The length of the pause, $\tau$, is a random variable; in the experiment of Cardoso *et al.*the pause is a trapping event in which a molecule is sequestered in the core of a stationary vortex. If the average duration of a pause is well defined then one can simply use Einstein's formula (6) with $\tau$ replaced by the average time between steps. However, if the pausing times are very broadly distributed then the average duration of a pause may be infinite and consequently the dispersion is subdiffusive. We explore this in more details in the next section.

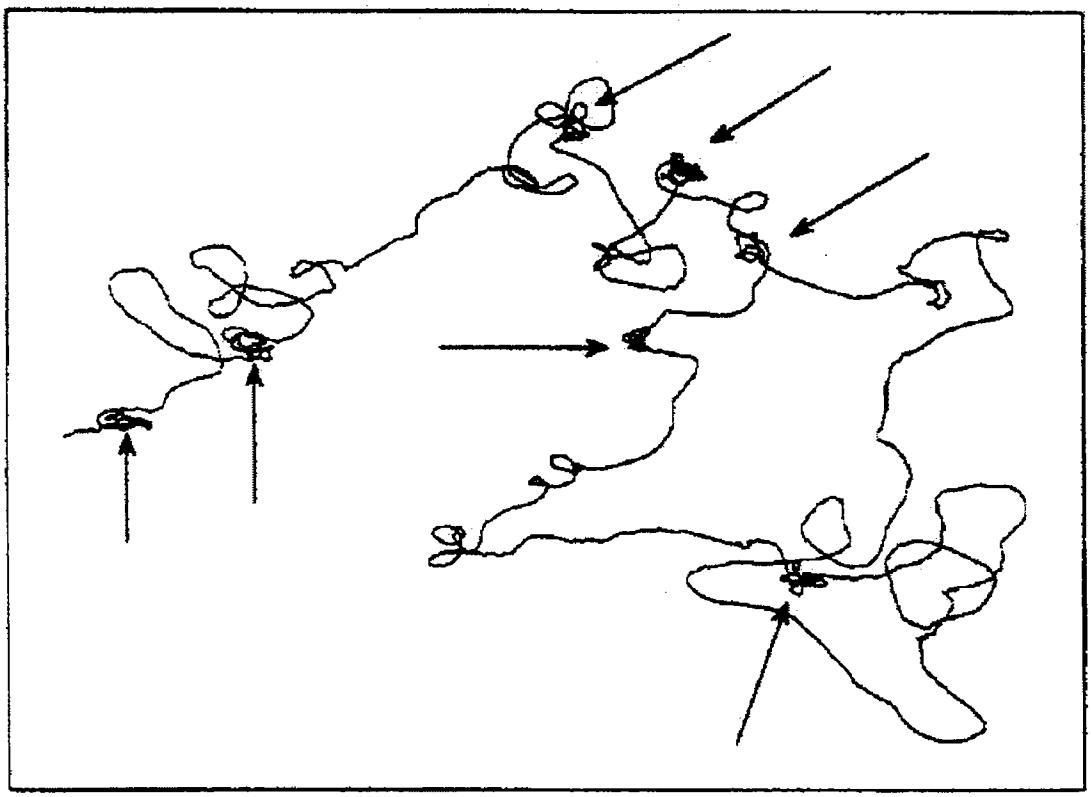

Figure 4: The trajectory of a single particle shows a sequence of long flights interrupted by trapping events in which the particle circles around a vortex. The vortex trapping events are indicated by the arrows. (From Cardoso *et al.* [2])

### 3.3 Random walk with pauses

Consider a random walk in which the walker pauses for a random time $\tau$ between steps. The various $\tau$'s have a probability density function $W(\tau)$ (the waiting time PDF). This PDF is normalised,

$$\int_0^\infty W(\tau)\,\mathrm{d}\tau = 1\,, \tag{28}$$

and the average waiting time spent between steps is

$$\bar{\tau} = \int_0^\infty \tau W(\tau)\,\mathrm{d}\tau\,. \tag{29}$$

Motivated by the experiments of Cardoso *et al.*, we entertain the notion that $\bar{\tau}$ is infinite because the integral in (29) diverges. For example, suppose that for large $\tau$, $W(\tau) \sim \tau^{-\mu}$. Then $\bar{\tau} = \infty$ if $\mu \leq 2$.

However, if we only observe a finite number of steps, then we do not sample the entire density $W(\tau)$. Specifically, suppose that after $N$ steps, we have experienced pauses of duration $\tau_1, \tau_2, \ldots, \tau_N$. We want to estimate the likely value of $\tau_{\max}(N) \equiv \max\{\tau_1, \tau_2, \ldots, \tau_N\}$. The quantity $\tau_{\max}(N)$ is useful because we can argue that the structure of $W(\tau)$ for $\tau > \tau_{\max}(N)$ cannot be significant for the displacement after $N$ steps.

To determine $\tau_{\max}(N)$, we turn to probability theory. Consider a random variable $\theta$ uniformly distributed in the interval $[0,1]$. That is, the PDF of $\theta$ is $P(\theta) = 1$ if $0 < \theta < 1$ and $P(\theta) = 0$ otherwise. Suppose we take $N$ samples, $\theta_1, \cdots, \theta_N$ and define $\theta_{\min}(N) \equiv \min\{\theta_1, \cdots, \theta_N\}$. In this simple case it is plausible that $\theta_{\min} \sim N^{-1}$ as $N \to \infty$.

Now the trick is to use $\theta$ to represent $\tau$: we write $\theta = \tau^p$, and adjust $p$ so that the power–law tail of $W(\tau) \sim \tau^{-\mu}$ corresponds to the simple structure of $P(\theta) = 1$. In fact,

$$P(\theta) = W(\tau)\left|\frac{\mathrm{d}\tau}{\mathrm{d}\theta}\right|, \qquad \Longrightarrow \qquad 1 \sim \tau^{1-\mu-p}, \tag{30}$$

or $p = 1 - \mu$. Because the minimum value of $\theta$ maps to the maximum value of $\tau$, it follows that

$$\tau_{\max}(N) \sim N^{1/(\mu-1)}. \tag{31}$$

Now we return to (29) to estimate the effective average pause time after $N$ pauses:

$$\bar{\tau}_{\text{eff}} = \int_0^{\tau_{\max}} \tau W(\tau)\,\mathrm{d}\tau \sim \tau_{\max}^{2-\mu}. \tag{32}$$

It is also plausible that the total time $t$ spent on this random walk is given by

$$t \sim N\bar{\tau}_{\text{eff}}. \tag{33}$$

Combining (31), (32) and (33) yields the following scaling relationships:

$$N \sim t^{\mu-1}, \qquad \bar{\tau}_{\text{eff}} \sim t^{2-\mu}, \qquad \tau_{\max} \sim t. \tag{34}$$

The final relation is worthy of comment: it implies a form of self-similarity of the random walk.

To conclude, the total displacement of our random walk is proportional to $\sqrt{N}$. But, with the random pauses, the scaling against time has been altered to

$$\text{RMS displacement} \propto \sqrt{N} \sim t^{(\mu-1)/2}. \tag{35}$$

This theory can be used to interpret the experiment of Cardoso *et al.*: because the RMS displacement grows as $t^{1/3}$ it follows that $\mu \approx 5/3$. Cardoso *et al.*successfully tested this prediction by measuring the PDF of trapping times inside vortices.

# 4 Stirring and mixing

## 4.1 Coffee and cream

Appealing to the everyday experience of mixing cream into coffee, Eckart [5] argued that the homogenization of two fluids occurs in three stages. The distinction between the stages is the value of the concentration gradient averaged over the domain.

***Initial***: there are distinct interfaces separating globules of cream and coffee. Within each globule, the concentration of cream is nearly constant and the concentration gradient is close to zero. There is a very large concentration gradient between regions of coffee and cream. But the interfaces between coffee and cream are small in number and not of great area, so the average gradient in the coffee mug is small.

***Stirring***: the cream is mechanically swirled and folded, and molecular diffusion is unimportant. During this second stage the concentration gradients increase.

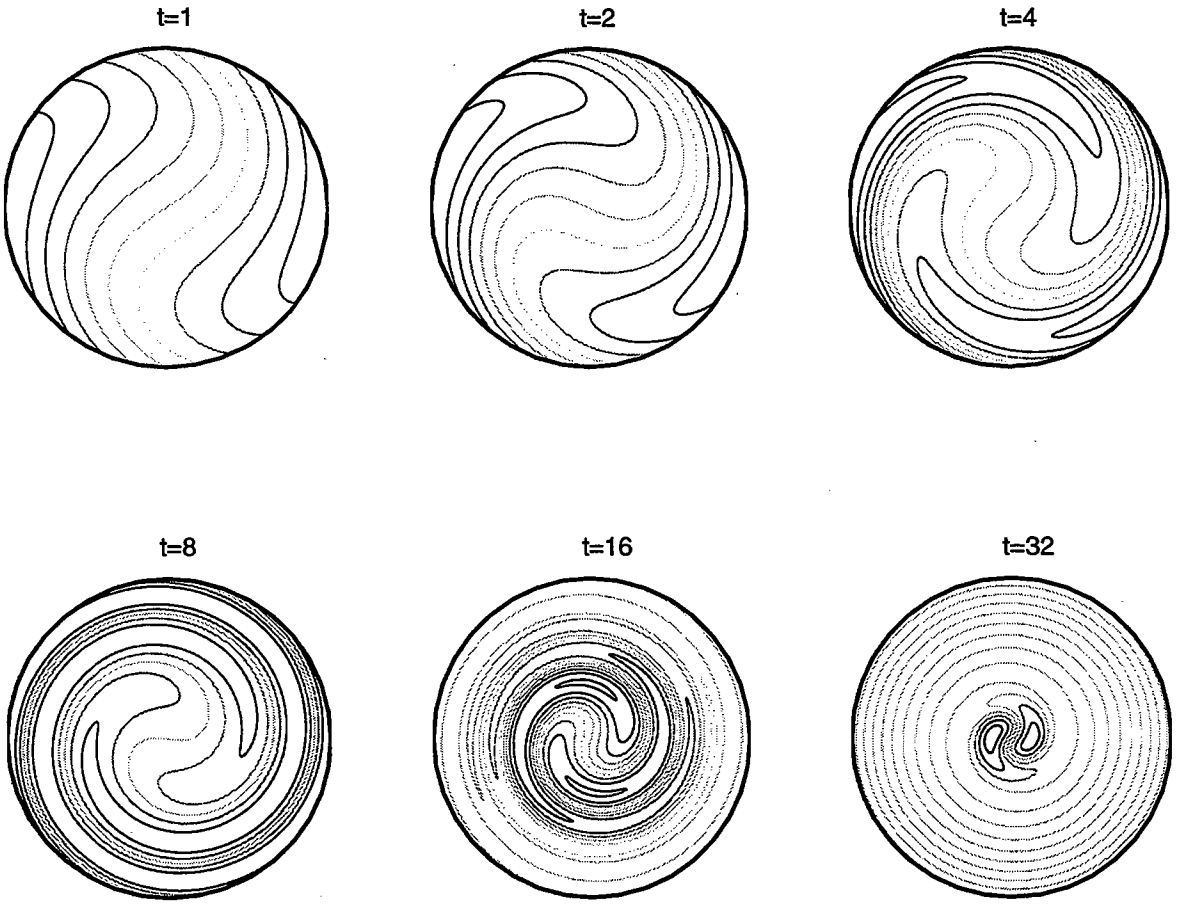


Figure 5: Solution of $c_t + (1-r^2)c_\theta = (8 \times 10^{-4})\nabla^2 c$. The initial condition is $c(x, y, 0) = x$.

***Mixing***: the gradients suddenly disappear and the fluid becomes homogeneous; molecular diffusion is responsible for the sudden mixing.

In a chemical reaction, molecules of different species must come into contact for the reaction to occur. Thus, when the species are initially separated, the reaction will not begin until the final mixing stage is reached. In this sense there is an important distinction between coarse-grained homogenization, occuring solely as a result of stirring, and mixing at the molecular scale.

To illustrate these concepts figure 5 shows a solution of the advection diffusion equation

$$c_t + (1-r^2)c_\theta = \kappa \nabla^2 c, \qquad c(r, \theta, 0) = r \sin\theta \tag{36}$$

where $\kappa = 8 \times 10^{-4}$. A particle at a distance $r$ from the origin completes a rotation in a time $2\pi/(1-r^2)$. Thus particles at smaller values of $r$ will overtake particles at larger values of $r$ and so the concentration is twisted into spirals by differential advection (stirring).

The increase in gradient during the stirring phase is evident in the figure. But at approximately $t = 16$, mixing starts to dominate, and diffusion rapidly reduces the average gradient. From the initial condition, an estimate of the time it would take unassisted diffusion to homogenize the fluid is $T_D \sim 1/\kappa = 1250$. It is only through the initial process of stirring that the concentration gradient is amplified or, alternatively, that the spirals are stretched out so that small diffusion homogenizes the tracer at $t = 32 \ll T_D$.

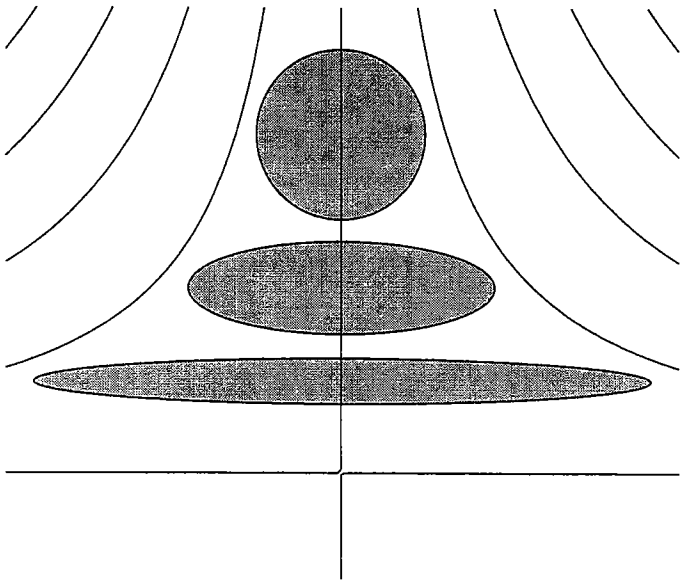

Figure 6: The straining flow described by the streamfunction $\psi = -\alpha xy$. The figure shows how a circular patch of tracer is stretched out along the $x$–axis by the hyperbolic strain. If $\kappa = 0$ the major axis of the ellipse grows as $\exp(\alpha t)$ and the minor axis reduces as $\exp(-\alpha t)$ so that the area remains constant.

## 4.2 A straining flow

A simple example of a two-dimensional flow which amplifies concentration gradients is the hyperbolic strain shown in figure 6. The streamfunction is $\psi = -\alpha xy$ and so the advection diffusion equation is

$$c_t + \alpha x c_x - \alpha y c_y = \kappa \nabla^2 c. \tag{37}$$

Notice the dimensions here: $\alpha^{-1}$ has dimensions "time" and $\kappa$ has dimensions (length)$^2$/(time). From these two quantities we can build a combination with the dimensions of (length):

$$\ell \equiv \sqrt{\frac{\kappa}{\alpha}}. \tag{38}$$

The length $\ell$ will appear prominently in the sequel.

We begin our discussion of hyperbolic strain by obtaining a solution in which $c$ is independent of both $x$ and $t$. In this special case the solution of (37) is

$$c_y = A \exp\left[-\frac{y^2}{2\ell^2}\right], \quad c(x, \pm\infty, t) = \pm\sqrt{2\pi} A \ell. \tag{39}$$

The concentration profile is the error function shown in figure 7. The solution shows the steady state balance between advection and diffusion: with $\sqrt{2\pi} A\ell = 1$, the concentration $c$ changes smoothly between $c = +1$ as $y \to +\infty$ to $c = -1$ as $y \to -\infty$. The transition occurs in a front of width $\ell$.

We can give an intuitive discussion of how the steady state profile in figure 7 is established as the solution of an initial value problem. Suppose we had started with the initial condition such as $c(x, y, 0) = \mathrm{sgn}(y)$ in which the transition between $c = -1$ and $c = +1$ occurs in a distance much less than $\ell$. Then the discontinuity in $c$ initially diffuses freely, growing like $\sqrt{\kappa t}$. Once the width of the front becomes comparable to $\ell$, that is when

$$\sqrt{\kappa t} \sim \ell, \qquad \Longrightarrow \qquad t \sim \alpha^{-1}, \tag{40}$$

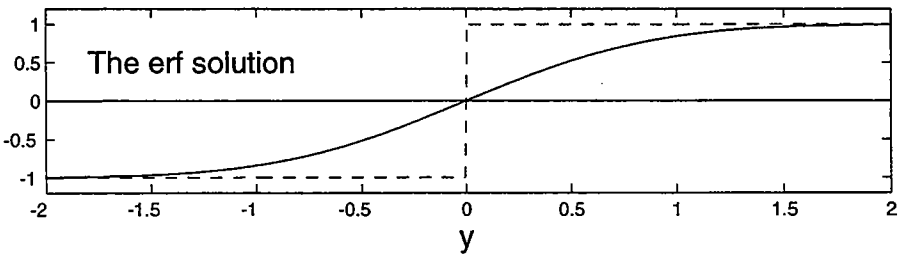


Figure 7: The time independent error function solution to equation (37).

the spread is arrested and the steady state in figure 7 is established.

On the other hand, we can also consider an initial condition in which the transition between $c = 1$ and $c = -1$ occurs on a scale $L_0 \gg \ell$. In this case the front is initially compressed by the hyperbolic strain so that the width is reduced exponentially, $L = L_0 \exp(-\alpha t)$. Because $L_0 \gg \ell$ the diffusion is unimportant until the exponential reduction in scale reaches $\ell$. That is,

$$L_0 e^{-\alpha t} \sim \ell\,, \qquad \Longrightarrow \qquad t \sim \alpha^{-1} \ln(L_0/\ell). \tag{41}$$

These considerations illustrate the fundamental importance of $\ell$ as the scale on which advection and diffusion come into balance.

## 4.3 Lagrangian coordinates: a simple example

The hyperbolic strain also provides a painless illustration of some mathematical techniques which can be used in more complicated problems. We begin by considering the solution of (37) with $\kappa = 0$. With no diffusion $c$ is tied to fluid particles. The position of a fluid particle is related to its initial position $(a, b)$, by solving the differential equations

$$(\dot{x}, \dot{y}) = \alpha(x, -y), \qquad \Longrightarrow \qquad (x, y) = \left(e^{\alpha t} a, e^{-\alpha t} b\right). \tag{42}$$

The solution of (37) can now be obtained by arguing that the particle which is at the point $(x, y)$ at time $t$ began at $(a, b) = (\exp(-\alpha t)x, \exp(\alpha t)y)$ at $t = 0$. Because the a particle carries the concentration it follows that the solution of (37) as an initial value problem is

$$c(x, y, t) = c_0\left[\exp(-\alpha t)x, \exp(\alpha t)y\right], \tag{43}$$

where $c_0(x, y)$ is the initial condition. The philosophy of this method is that we care where fluid particles come from, but not where they are going to.

The solution above seems to rely crucially on the restriction that $\kappa = 0$. But now look what happens if we use the Lagrangian coordinates $(a, b)$ in (42) as new independent variables in (37). As an accounting device, it is comforting to define $\tau = t$ and consider that $\partial_\tau$ as the time derivative with $(a, b)$ fixed. Thus the transformation rules are

$$(\partial_x, \partial_y) = \left(\frac{\partial a}{\partial x}, \frac{\partial a}{\partial y}\right)\partial_a + \left(\frac{\partial b}{\partial x}, \frac{\partial b}{\partial y}\right)\partial_b = \left(e^{-\alpha\tau}\partial_a, e^{\alpha\tau}\partial_b\right). \tag{44}$$

and

$$\partial_t = \frac{\partial \tau}{\partial t}\partial_\tau + \frac{\partial a}{\partial t}\partial_a + \frac{\partial b}{\partial t}\partial_b = \partial_\tau - \alpha a \partial_a + \alpha b \partial_b. \tag{45}$$

The punchline is that

$$\partial_t + \alpha x \partial_x - \alpha y \partial_y = \partial_\tau\,, \tag{46}$$

which shows that the change to a Lagrangian description makes the convective derivative trivial.

Substituting the transformations above into (37) gives:

$$c_t = \kappa e^{-2\alpha t} c_{aa} + \kappa e^{2\alpha t} c_{bb}. \tag{47}$$

Naturally, if $\kappa = 0$, we recover our earlier solution in (43). But even if $\kappa \neq 0$ it is often easier to solve (47) than the Eulerian form in (37). For example, Fourier transforming (47), with $(\partial_a, \partial_b) \to \mathrm{i}(p, q)$, gives a simple ordinary differential equation in time.

It is instructive to use the method above to solve (37) with the initial condition

$$c(x, y, 0) = \delta(x)\delta(y)\,. \tag{48}$$

Physically, this is a spot of dye released in a straining flow. When $\alpha t \ll 1$ the spot spreads diffusively, with a diameter which grows as $\sqrt{\kappa t}$. However when $\alpha t \sim 1$ the diameter of the spot becomes comparable to $\ell \equiv \sqrt{\kappa/\alpha}$, and then the spot stops expanding against the compressive direction of the strain. However the spot continues to stretch along the extensive direction. Thus, when $\alpha t > 1$, the spot becomes a filament with an equilibrium width of order $\ell$ and an exponentially growing length. These intuitive arguments are supported by the exact solution:

$$c(x, y, t) = \frac{1}{4\pi f g} \exp\left[-\frac{x^2}{4f^2} - \frac{y^2}{4g^2}\right], \tag{49}$$

where $f(t)$ and $g(t)$ are

$$f^2 \equiv \frac{\kappa}{2\alpha}(e^{2\alpha t} - 1)\,, \qquad g^2 \equiv \frac{\kappa}{2\alpha}\left(1 - e^{-2\alpha t}\right)\,. \tag{50}$$

Notice that the peak concentration ultimately decreases like $e^{-2\alpha t}$.

## 4.4 An example of sudden mixing

As a final look at the hyperbolic straining flow, we note that a solution of (37) is

$$c(x, y, t) = A(t)\cos(k e^{-\alpha t} x)\cos(k e^{\alpha t} y), \tag{51}$$

where

$$A(t) = \exp\left[-\ell^2 k^2 \sinh 2\alpha t\right]\,. \tag{52}$$

One route to this exact solution is to look for separable solutions of (47), and then transform back to the Eulerian coordinates (e.g., Young, Rhines & Garrett,1982).

The mean value of the square of the concentration gradient varies with time as:

$$\{\boldsymbol{\nabla} c \cdot \boldsymbol{\nabla} c\} = \frac{k^2}{2}\cosh(2\alpha t)\exp\left[-2\ell^2 k^2 \sinh(2\alpha t)\right]\,, \tag{53}$$

where $\{\}$ denotes an average over a large area. $\{\boldsymbol{\nabla} c \cdot \boldsymbol{\nabla} c\}$ is plotted in figure 8 for various values of $k\ell$. Recalling Eckart's description of stirring as increasing the concentration gradient, and mixing as decreasing the concentration gradient, we can see the transition between the two phases occurs at the peaks of the various curves. If $k\ell \ll 1$, then the time it takes to reach this peak is given by $t_*$, where

$$\alpha t_* \sim -\ln(k\ell). \tag{54}$$

Once again, this is the time taken for the exponential factor $e^{-\alpha t}$ to reduce initial length of the tracer field, $k^{-1}$, down to the length $\ell$ on which strain and diffusion balance.

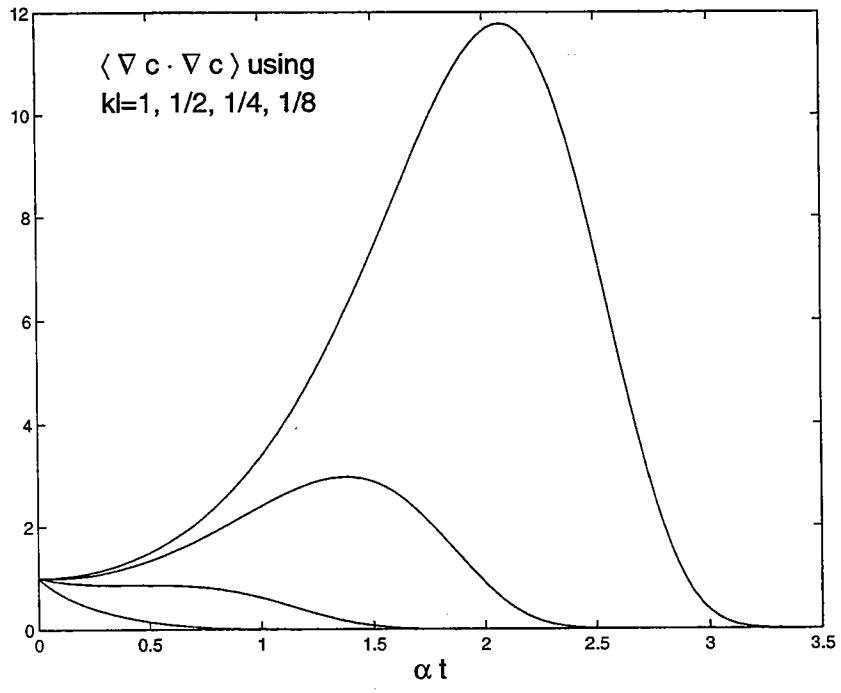


Figure 8: The mean square of the concentration gradient. If $k\ell \ll 1$ then the concentration gradient grows until $t = t_*$ in (54) and then decreases precipitiously. If $k\ell \geq 1/2$ then diffusion always overpowers strain and the mean square gradient decreases monotonically to zero.

## 4.5 A Welander scrapbook

Stirring was beautifully illustrated in a 1955 paper of Welander's [12]. This paper is notable also because of its discussion of the importance of coarse-grained averages. Figures 9, 10 and 11 reproduced from Welander (1955) show that simple velocity fields produce spectacular distortion of passive scalars.

In figures 9, 10 and 11, some dimensions of the scalar blob are stretched out while other dimensions are contracted. Batchelor (1952) [1] argued that in turbulent flows random stretching results in an exponential growth of the separation between two initially adjacent fluid elements. That is, if we consider two material elements separated by a distance $s_0$ which is much less than the scale of the velocity field, then Batchelor argues that the separation grows as

$$s \sim s_0 e^{\gamma t} . \tag{55}$$

The time-scale $\gamma^{-1}$ is analogous to $\alpha^{-1}$ in (37), though in figures 10 and 11 the exponential straining is driven by a random and unsteady velocity, rather than the simple hyperbolic field in figure 6. Note particularly that the exponential law in (55) is valid until the separation $s(t)$ becomes comparable to the length scale over which the velocity varies.

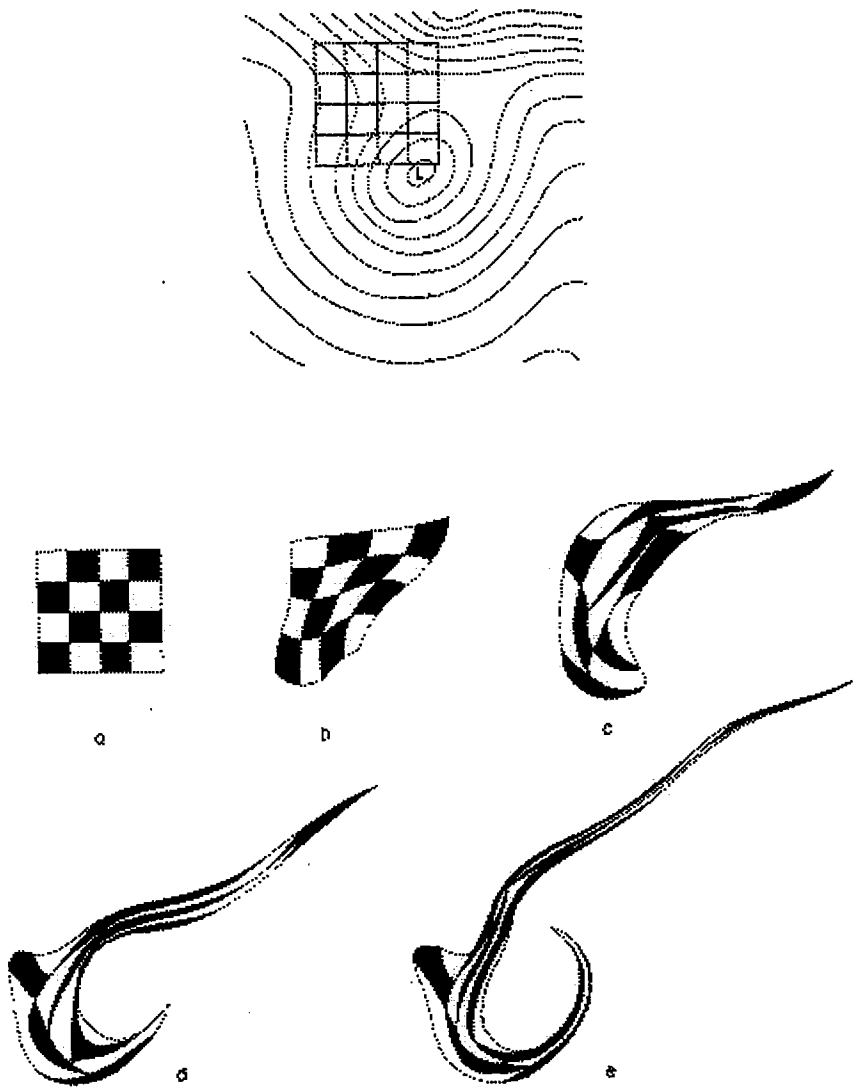


Figure 9: Welander's numerical solution illustrating differential advection by a simple velocity field. A checkerboard pattern is deformed by a quasigeostrophic barotropic solution which models atmospheric flow at the 500mb level. The initial streamline pattern is shown at the top and the subsequent figures are at 6 hours, 12 hours, 24 hours and 36 hours, respectively. Notice that each square of the checkerboard maintains constant area as it deforms.

# References

[1] G.K. Batchelor. The effect of turbulence on material lines and surfaces. *Proc. Roy. Soc. London A*, 213:349–366, 1952.

[2] O. Cardoso, B. Gluckmann, O. Parcollet, and P. Tabeling. Dispersion in a quasi-two-dimensional turbulent flow: an experimental study. *Phys. Fluids*, 8(1):209–214, January 1996.

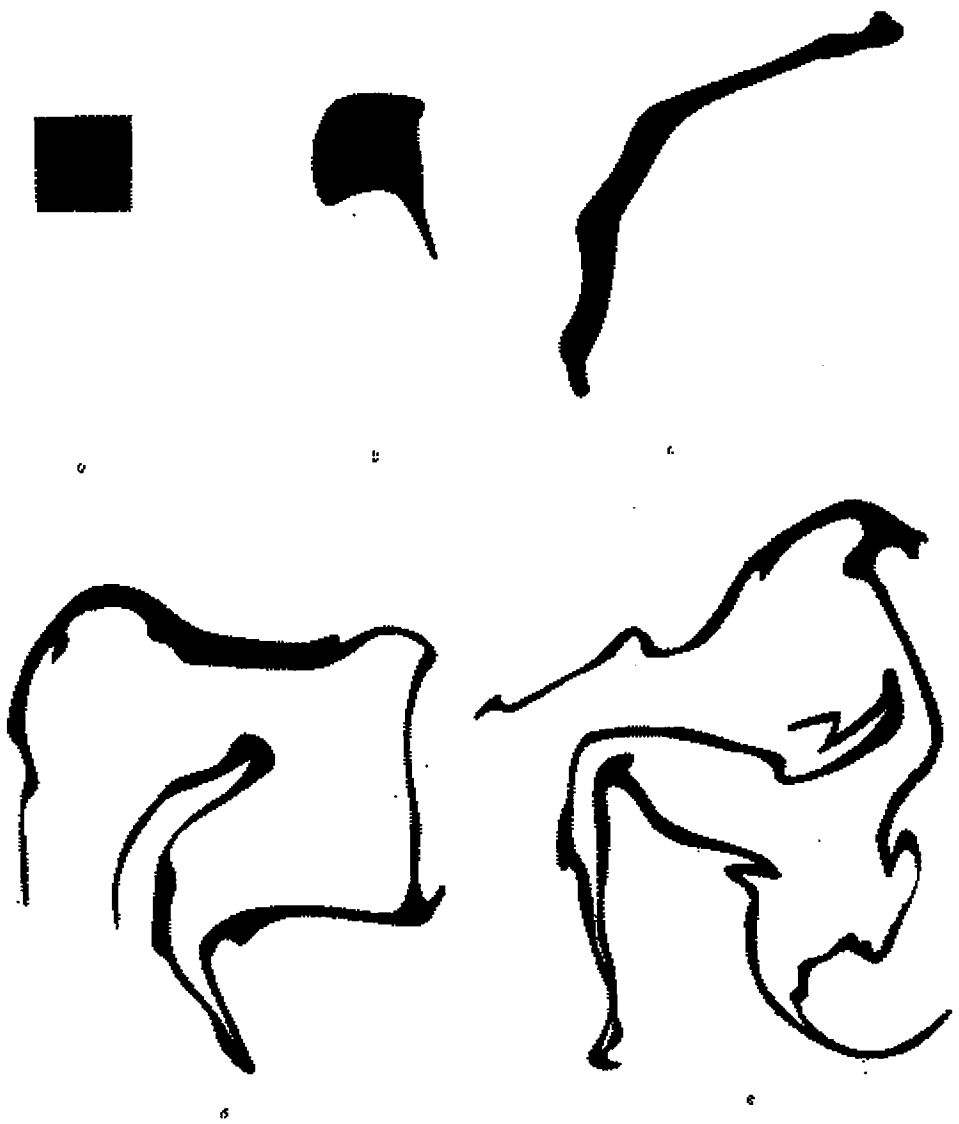

Figure 10: Welander's experimental illustration of the deformation of a small coloured square element of a fluid surface. To suppress three dimensional turbulence, a vessel of water is brought into solid body rotation. A floating film of butanol is divided into square elements by means of a metal grid and one of these square elements is coloured with methyl-red. The solid body rotation is disturbed by stirring the water and the grid is removed.


[3] O. Cardoso and P. Tabeling. Anomalous diffusion in a linear array of vortices. *Europhys. Lett.*, 7(3):225–230, 1988.

[4] O. Cardoso and P. Tabeling. Anomalous diffusion in a linear system of vortices. *Euro. J. Mech. B/Fluids*, 8(6):459–470, 1989.

[5] C. Eckart. An analysis of stirring and mixing processes in incompressible fluids. *J. Mar. Res.*, 7:265–275, January 1948.

Figure 11: Further deformation of the butanol square in figure 10. According to Batchelor [1], the length of the filaments increases exponentially with $t$.


[6] A. Einstein. *Investigations on the theory of the Brownian Movement.* Dover, New York, 1956.

[7] A. Pais. *'Subtle is the Lord...'.* Oxford University Press, Oxford and New York, 1982.

[8] T.H. Solomon and J.P. Gollub. Chaotic particle transport in time dependent Rayleigh-Bénard convection. *Phys. Rev. A*, 38:6280, 1988.

[9] T.H. Solomon and J.P. Gollub. Passive transport in steady Rayleigh-Bénard convection. *Phys. Fluids A*, 31:1372–???, 1988.

[10] T.H. Solomon, S. Tomas, and J.L. Warner. Chaotic mixing of immiscible impurities in a two-dimensional flow. *Phys. Fluids A*, 10:342–350, 1998.

[11] G.I. Taylor. Diffusion by continuous movements. *Proc. London Math. Soc.*, 20:196–212, 1921.

[12] P. Welander. Studies on the general development of motion in a two-dimensional, ideal fluid. *Tellus*, 7:141–156, 1955.

[13] W.R. Young, A. Pumir, and Y. Pomeau. Anomalous diffusion of tracer in convection rolls. *Phys. Fluids A*, 1:462–469, 1989.

[14] W.R. Young, P.B. Rhines, and C. Garrett. Shear-flow dispersion, internal waves and horizontal mixing in the ocean. *J. Phys. Ocean.*, 12:515–527, 1982.

# Lecture 2: Diffusion

## 1 Introduction

Perhaps you have heard that turbulence is the most difficult problem in fluid mechanics and, according to some, the greatest unsolved problem in physics. One indication of the difficulty is that it is impossible to give a satisfactory definition of a "turbulent flow". But everyone agrees that one property of turbulence is greatly enhanced transport of passive contaminants. For example, relying only on molecular agitation, a dissolved sugar molecule takes years to diffuse across a coffee cup, and on that time-scale the coffee will surely evaporate. With a spoon the coffee drinker can create eddies that transport dissolved sugar throughout the cup in less than one second. This is an example of *eddy diffusivity.*

Fluid mechanics textbooks often often justify eddy diffusivity by appealing to an analogy between turbulent eddies and molecular diffusion — perhaps this notion originates with G.I. Taylor's 1905 paper entitled "Eddy motion in the atmosphere" [4]. In any event, the molecular analogy, supplemented with some hand-waving, leads to the notion of an eddy diffusivity and for many scientists this is the end of the turbulence problem.

Our goal in this lecture is to explain very explicitly the assumptions behind Taylor's "proof by analogy" and to illustrate the interesting points at which the analogy fails. We will pursue this program by working with some very simple model flows for which analytic results, such as expressions for the eddy diffusivity, are available. As you will soon see, these model flows do not greatly resemble turbulence, but then neither does molecular motion! Our excuse is that soluble examples are always diverting and educational.

## 2 The renovating wave model

### 2.1 A recipe for constructing soluble models

The main problem in analyzing transport is solving the differential equations which describe the motion of particles in even very simple flows. However there is a class of flows for which this task is trivial. These are steady and unidirectional flows, such as $u = \sin y$. A particle which starts at $(a, b)$ at $t = 0$ finds itself at $(a + \tau \sin b,\, b)$ at $t = \tau$. This is dull, but it becomes more interesting if at intervals of $\tau$ we "renovate" the flow by randomly picking a new direction along which the velocity acts. In this way we can construct a sequence of iterated random maps and calculate diffusivities, and other statistical properties, by averaging the exact solution. I learned of this trick from the literature on dynamo theory. The book *Stretch, Twist, Fold: the Fast Dynamo* is highly recommended for students interested in all aspects of stirring and mixing [1].

### 2.2 The renovating wave (RW) model

As a particular example we now formulate the *renovating wave* (RW) model. We divide the time axis into intervals

$$I_n \equiv \{t : (n-1)\tau < t < n\tau\}\,, \tag{1}$$

and in each interval we apply a velocity, $\mathbf{u} = (-\psi_y, \psi_x)$, derived from the streamfunction

$$\psi_n(x, y, t) = k^{-1} U \cos[k \cos\theta_n \, x + k \sin\theta_n \, y + \varphi_n] \,, \tag{2}$$

where $\theta_n$ and $\varphi_n$ are independent random variables uniformly distributed in the interval $[-\pi, \pi]$. Thus in each $I_n$ there is a steady, unidirectional velocity with sinusoidal profile (a single wave). There is sudden and complete loss of all information about the past velocity at $t = n\tau$ because at these instants we "renovate" the velocity by picking new random angles $\theta$ and $\varphi$. (This means that the velocity correlation function, $\mathcal{C}(t)$, is zero if $t > \tau$.)

The renovating wave model can be nondimensionalized by using $k^{-1}$ as a unit of length and and $1/(Uk)$ as a unit of time. With this choice, the model contains a single dimensionless parameter, $\tau_* \equiv \tau k U$, which is a measure of the persistence of the motion. Much of the literature on random advection-diffusion uses model velocity fields which are $\delta$-correlated in time. We can recover this limit as a special case by taking $\tau_* \to 0$.

Using dimensionless variables, a particle which is at $\mathbf{x}_n = (x_n, y_n)$ at $t_n = n\tau_*$ moves to $\mathbf{x}_{n+1}$ at $t = (n+1)\tau_*$, where

$$(x_{n+1}, y_{n+1}) = (x_n, y_n) + \tau_* \sin(c_n x + s_n y + \varphi_n)\,(s_n, -c_n) \,. \tag{3}$$

with $s_n \equiv \sin\theta_n$ and $c_n \equiv \cos\theta_n$. Thus motion in the renovating wave problem is equivalent to an iterated sequence of random maps.

## 2.3 The single–particle diffusivity

It is very easy to calculate the diffusivity in the RW model (and much more difficult to interpret the answer). The average of a function of the two random angles $\theta$ and $\varphi$ (suppress the subscript $n$) is defined by

$$\langle f \rangle = \oint \frac{\mathrm{d}\varphi}{2\pi} \oint \frac{\mathrm{d}\theta}{2\pi} f(\theta, \varphi). \tag{4}$$

Therefore, using (3),

$$\langle (x_{n+1} - x_n)^2 \rangle = \frac{\tau_*^2}{4} \,. \tag{5}$$

The computation is trivial if the integral over $\varphi$ is evaluated first.

In (5), following our previous discussion based on Einstein's derivation of the diffusion equation, we are computing the statistics of dispersion along the $x$-axis. Because the renovating wave model is isotropic, dispersion in the $y$-direction is identical to that in the $x$-direction.

Because all of the waves are independent and identically distributed it follows that after $n$ renovation cycles

$$\langle (x_n - x_0)^2 \rangle = n \frac{\tau_*^2}{4}. \tag{6}$$

But $t = n\tau_*$, and $\langle (x_n - x_0)^2 \rangle = 2Dt$, so that using dimensionless variables the diffusivity is

$$D = \frac{\tau_*}{8}. \tag{7}$$

Sometimes $D$ is referred to as the single-particle diffusivity. "Single-particle" emphasizes that $D$ strictly applies only to the RMS displacement of a particle from its initial position; $D$ contains no information concerning the deformation of a patch of tracer, nor of any other quantity involving correlated motion. Thus, using dimensional variables, the diffusivity in (7) is $D = U^2\tau/8$, which is independent of $k$. Because $D$ is independent of the scale of the wave, even a spatially uniform, but random-in-time velocity (the case $k = 0$), has a single-particle diffusivity.

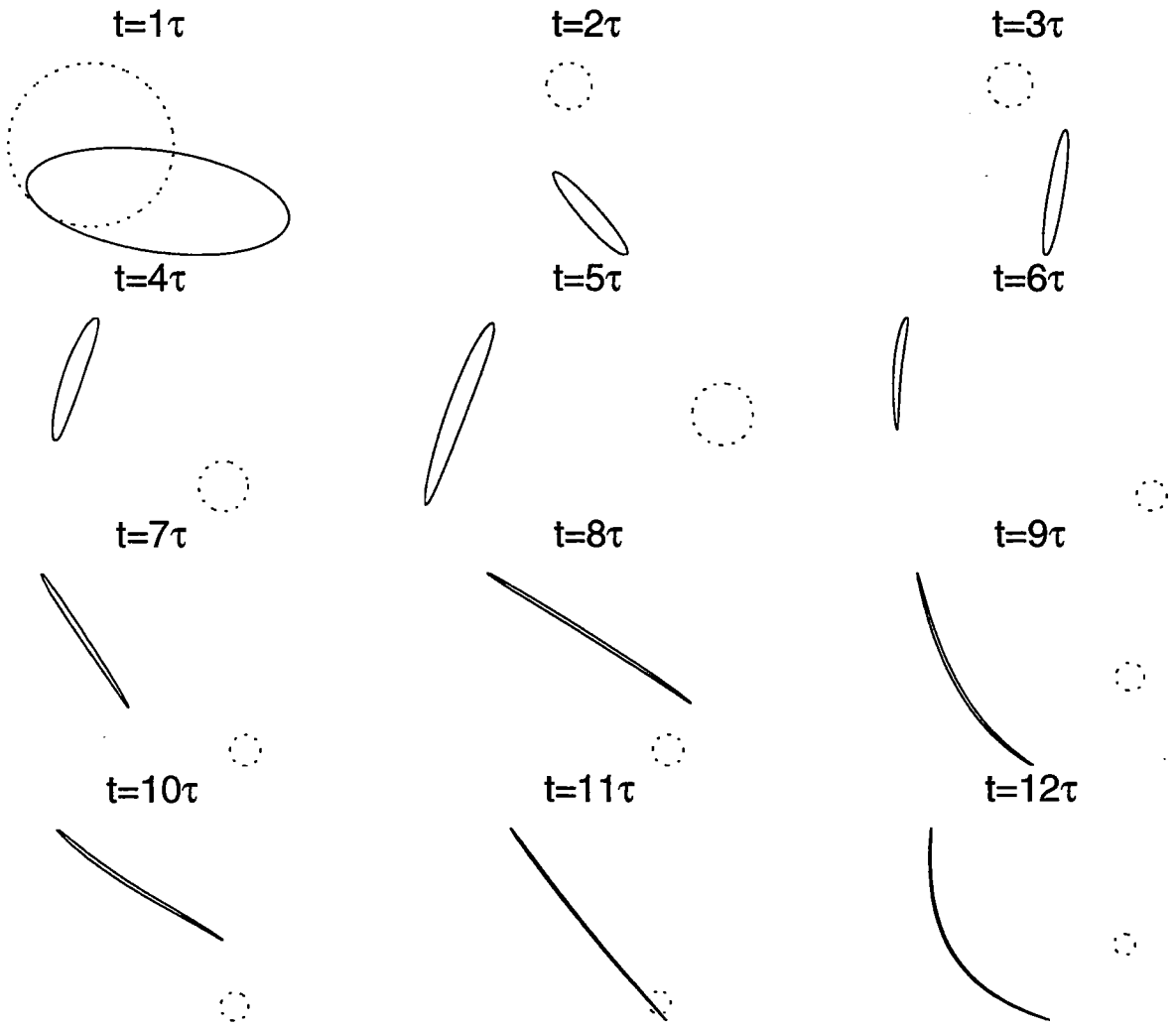


Figure 1: Stretching of a small spot, $r \ll 1$ where $r$ is the initial radius of the spot, by a succession of random sinusoidal flows. The dotted circle is the initial spot.

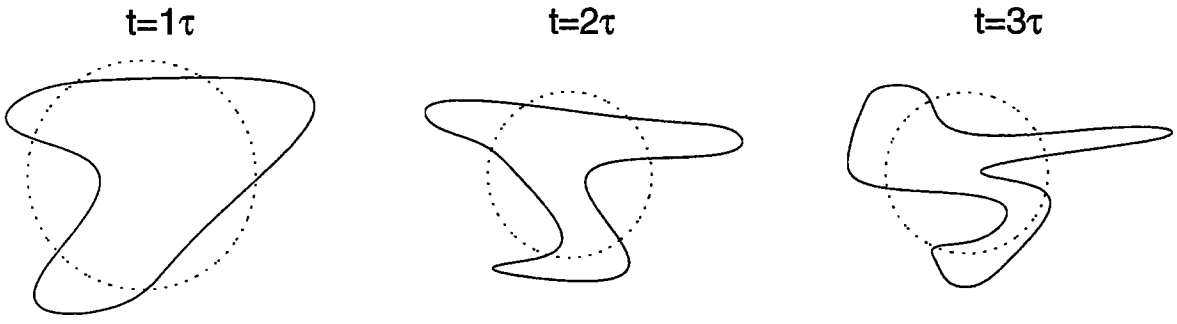


Figure 2: Stretching of a blob with $r = 1$, where $r$ is the initial radius. The dotted circle is the initial patch.

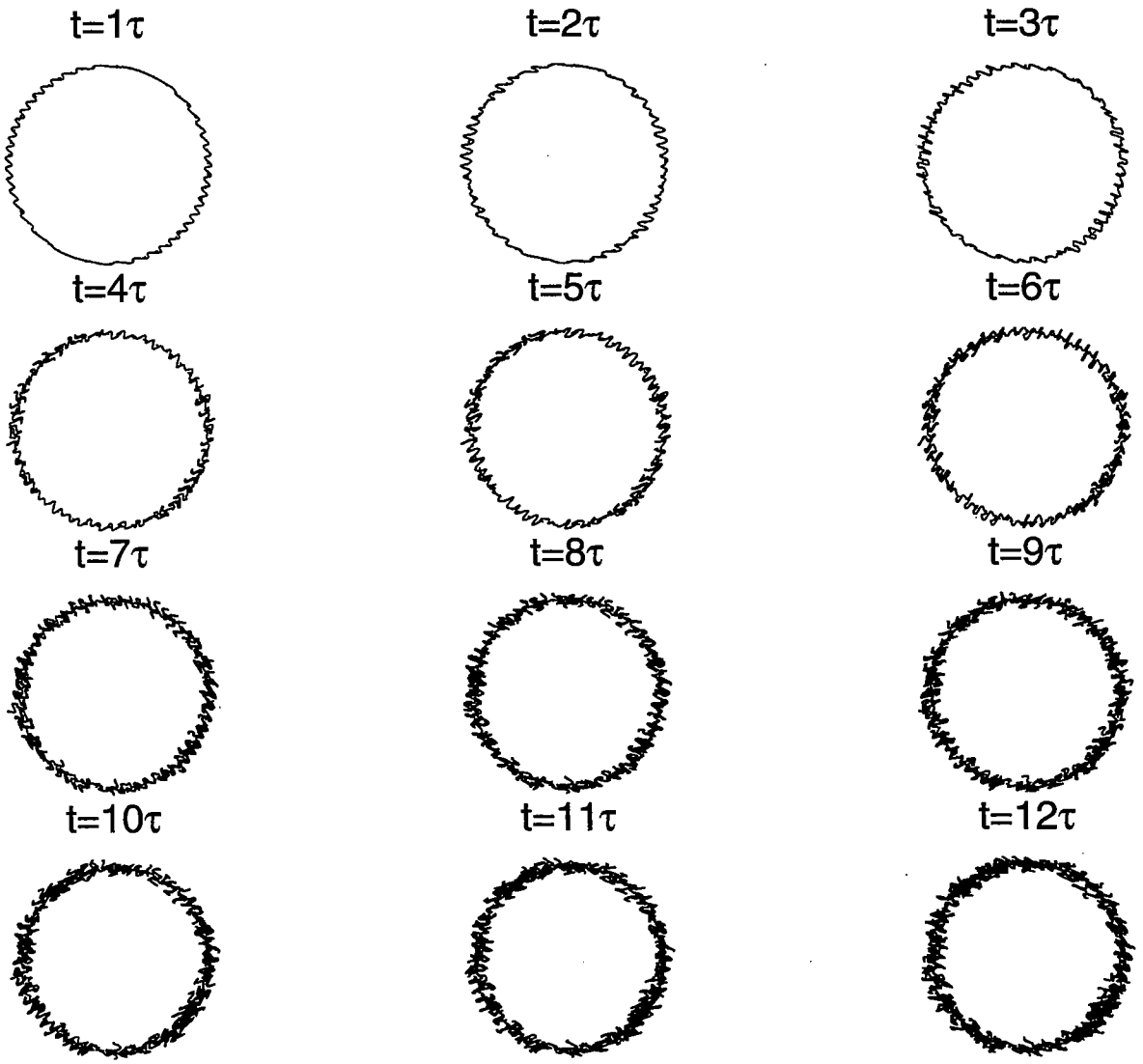


Figure 3: Stretching of a big blob $r \gg 1$, where $r$ is the initial radius of the blob. The dotted circle representing the initial patch may not be visible beneath the wiggly boundary of the blob.

## 2.4 Deformation of variously sized blobs

To emphasize the importance of understanding more than single-particle diffusivities we take a digression and illustrate how the deformation of an initially circular blob of fluid depends on the blob radius $r$. (Recall that we have used $k^{-1}$ as unit of length; in terms of dimensional variables the relevant nondimensional parameter is $kr$.)

If the initial blob is much smaller than the wavelength of the velocity then on the scale of the blob the velocity profile is a linear function of the coordinates. Because of this simplicity, the first few iterations deform the circular blob into an ellipse which must have the same area as the initial circle. We will see in the next lecture that the major axis of the ellipse grows exponentially while the minor axis shrinks so that the area is fixed. Once the dimensions of the ellipse are comparable to the wavenumber of the flow, more complicated deformations occur. Ultimately the blob will be stretched

into a folded filament as in figure 1.

The blob has the same scale as the velocity field if $r \sim 1$. Because there is no *scale separation* there is no easy description of the action of the flow on the blob, see figure 2.

If $r \gg 1$ then we are in the "eddy diffusivity" limit in which the scale of the velocity field is much smaller than the scale of the tracer. This case is shown in figure 3. The action of the waves perturbs the edge of the blob, making it look "fuzzy". In fact, the area is preserved, but the circumference of the blob grows exponentially. We will be discussing this type of problem for the remainder of the lecture.

### 2.5 The Lagrangian correlation function

In (7) we gave the diffusivity of particles moving in an ensemble of renovating waves. How do we obtain the Lagrangian velocity autocorrelation function and verify Taylor's formula that

$$D = \int_0^\infty C(t)\,\mathrm{d}t\,? \tag{8}$$

Considering this question, we encounter an annoying technical difficulty: our derivation of (8) assumes that the velocity statistics are stationary. But the renovating wave ensemble, as we defined it back in (1) and (2), is not a stationary stochastic process. This is because with our original definition all members of the ensemble renovate at the same instants $t = \tau$, $t = 2\tau$ etcetera. In order to obtain a stationary process we should initiate different realizations at uniformly distributed points during the renovation cycle. Thus, for realization number $j$, we pick a random time $\tau^{(j)}$ which is uniformly distributed in the interval $[0, \tau]$ and renovate first at $t = \tau^{(j)}$ and then subsequently at $t = \tau^{(j)} + \tau$, $t = \tau^{(j)} + 2\tau$ etcetera. With this new and improved formulation of the RW model the Lagrangian correlation function of $u(t)$ is a "triangular" function:

$$C(t) = \frac{U^2}{4}\left(1 - \frac{t}{\tau}\right) H(\tau - t)\,, \tag{9}$$

where $H$ is the step function and $U$ is the velocity in (2). The area under this correlation function is $D = U^2\tau/8$.

## 3 The eddy diffusion equation

### 3.1 The ensemble averaged Green's function

Now that we have obtained the RW diffusivity in (7) we turn to the derivation of the eddy diffusion equation. For each realization we introduce the Green's function which is

$$G_t + \mathbf{u}\cdot\boldsymbol{\nabla} G = 0, \quad \text{with} \quad G(\mathbf{x}, \mathbf{x}_0, 0) = \delta(\mathbf{x} - \mathbf{x}_0). \tag{10}$$

The solution of the problem above is

$$G(\mathbf{x}, \mathbf{x}_0, t) = \delta(\mathbf{x}_t - \mathbf{x}_0)\,, \tag{11}$$

where $\mathbf{x}_t$ is the position at time $t$ (in a particular realization of $\mathbf{u}$) of the particle which started at $\mathbf{x}_0$.

The ensemble averaged Green's function is

$$g(r, t) = \langle G(\mathbf{x}, \mathbf{x}_0, t)\rangle\,, \qquad r \equiv |\mathbf{x} - \mathbf{x}_0|\,, \tag{12}$$

where we have assumed that the random velocity is isotropic, homogeneous and stationary so that $g$ can depend only on the distance $r$ and the elapsed time $t$.

Possessing $g(r,t)$, we can then represent the ensemble-averaged solution of the initial value problem

$$c_t + \mathbf{u}\cdot\boldsymbol{\nabla} c = 0, \qquad c(\mathbf{x},0) = c_0(\mathbf{x}), \tag{13}$$

as the convolution

$$\langle c\rangle(\mathbf{x},t) = \int c_0(\mathbf{x}-\mathbf{x}')g(|\mathbf{x}'|,t)\,\mathrm{d}\mathbf{x}'. \tag{14}$$

(We are assuming that the initial condition $c_0$ is the same for all realizations.)

At this point, the analogy between (14) and the master equation of lecture 1 is obvious. With the master equation in mind, we can anticipate that a variant of Einstein's derivation of the diffusion equation can be applied to (14). Rather than develop a general derivation we prefer to use the renovating wave model as a concrete illustration of how one can obtain $g$, and then pass from the integral equation in (14) to an approximate diffusion equation.

## 3.2 The averaged Green's function of the RW model

There are at least two ways of obtaining $g(r)$ in (12) for the RW model: the hard, straightforward way (see the appendix) and the easy, devious way. Let us be devious.

We begin by calculating the probability density function (PDF) of displacements in a single pulse of the RW model. Because the ensemble of velocities is isotropic and homogeneous there is no harm in supposing that the particle is at the origin and the $x$-axis is aligned with the direction of the velocity. That is, put $(x_n, y_n) = (0,0)$ and $\theta_n = \pi/2$ in (3). Thus, the displacement $r$ produced by a single pulse is

$$x_{n+1} - x_n = \tau_* \sin\varphi_n, \quad \text{and} \quad r = |x_{n+1} - x_n|. \tag{15}$$

The PDF of the random variable $r$ can be obtained from the PDF of $\varphi$, that is $P(\varphi) = 1/2\pi$, using the rule for transforming probabilities:

$$P(r) = \sum P(\varphi)\left|\frac{\mathrm{d}\varphi}{\mathrm{d}r}\right|, \quad \Longrightarrow \quad P(r) = \frac{2}{\pi}\frac{H(\tau_* - r)}{\sqrt{\tau_*^2 - r^2}}. \tag{16}$$

In (16) $H(\tau_* - r)$ is a Heaviside step function which ensures that there are no displacements greater than $\tau_*$. (The sum in (16) is because there are four values of $\varphi$ corresponding to a single value of $r$.)

The averaged Green's function is now given by

$$g(r) = \frac{P(r)}{2\pi r}, \qquad \Longrightarrow \qquad g(r) = \frac{1}{\pi^2}\frac{H(\tau_* - r)}{r\sqrt{\tau_*^2 - r^2}}. \tag{17}$$

The geometric factor $2\pi r$ is included because $g(r)$ is a concentration. That is, $P(r)\mathrm{d}r$ the expected number of particles which fall into the differential annulus between $r$ and $r + \mathrm{d}r$ and $g(r)$ is the expected number of particles *per area* in this same annulus; see figure 4.

Now that we have the averaged Green's function of a single pulse we can obtain the evolution the ensemble averaged concentration, $\langle c\rangle$, over many pulses. Because each pulse is independent of the preceeding pulses we have

$$\langle c\rangle\,(\mathbf{x}, (n+1)\,\tau_*) = \int \langle c\rangle(\mathbf{x}-\mathbf{x}', n\tau_*)g(|\mathbf{x}'|)\,\mathrm{d}\mathbf{x}'. \tag{18}$$

The master equation above, with $g(r)$ in (17), is an exact description of the evolution of $\langle c\rangle$ under advection by the RW model.

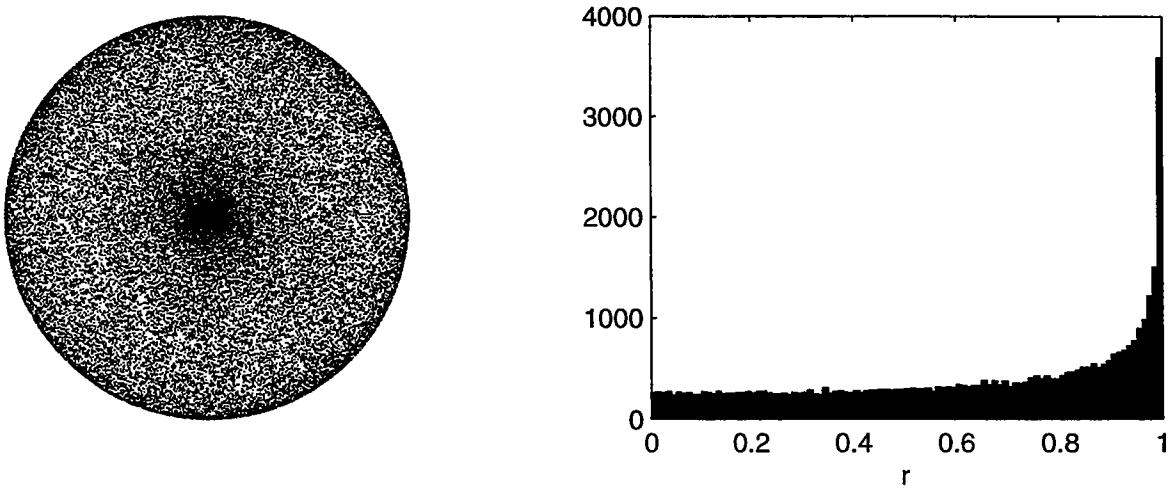


Figure 4: Displacements of 40,000 particles in independent realizations of the the RW model. The left panel shows the final position of particles which all start at the center of the the circle after one pulse of the wave. The density of points corresponds to $g(r)$ in (17). The histogram on the right shows the number of particles at a distance $r$ from the center; this is the function $P(r)$ in (16).

## 3.3 The diffusion equation

With the master equation (18) in hand, we can use Einstein's approximations to obtain the diffusion equation. Using the dimensionless variables of the renovating wave model, we have

$$\langle c\rangle_t \approx \frac{\tau_*}{8}\nabla^2\langle c\rangle\,. \tag{19}$$

We leave this as a homework exercise and instead we take a different route to (19).

Because the Fourier transform of a convolution is the product of the Fourier transforms, we can simplify (18) by transforming. The Fourier transform of $f(\mathbf{x})$ is defined here[1] as

$$\tilde{f}(\mathbf{k}) = \int e^{-i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{x})\,\mathrm{d}x\,, \qquad f(\mathbf{x}) = \frac{1}{2\pi}\int e^{i\mathbf{k}\cdot\mathbf{x}}\tilde{f}(\mathbf{k})\,\mathrm{d}k\,. \tag{20}$$

Applying the transform to (18) we obtain

$$\widetilde{\langle c\rangle}(\mathbf{k}, n\tau_*) = \tilde{g}(k)^n \tilde{c}_0(\mathbf{k})\,, \qquad k \equiv |\mathbf{k}|\,. \tag{21}$$

With a good table of integrals one can discover that the Fourier transform of the averaged Green's function, $g(r)$ in (17), is

$$\tilde{g}(k) = J_0^2(k\tau_*/2)\,, \tag{22}$$

where $J_0$ is the Bessel function.

The diffusion equation describes the evolution of large spatial scales, which is the same as small wavenumbers. This means that we simplify (21) by taking $k\tau_*/2 \ll 1$ and using the approximation $J_0(k\tau_*/2) \approx 1 - (k^2\tau_*^2/16)$ to write

$$\widetilde{\langle c\rangle}(\mathbf{k}, n\tau_*) \approx \exp\left\{n \ln\left[1 - (k^2\tau_*^2/8)\right]\right\} \tilde{c}_0(\mathbf{k})\,. \tag{23}$$

[1] By denoting the wavenumber with $k$ we are recycing notation used in (2).

But now, since $n = t/\tau_*$ and $\ln[1 - (k^2\tau_*^2/8)] \approx -k^2\tau_*^2/8$, we have

$$\widetilde{\langle c\rangle}(\mathbf{k}, t) = e^{-Dk^2t}\tilde{c}_0(\mathbf{k}) \tag{24}$$

where, as in (19), $D = \tau_*/8$. Equation (24) is the equivalent to the decay of Fourier components given by (19).

This derivation based on Fourier analysis explicitly recognizes that the diffusion approximation is valid only for wavenumbers which satisfy $k\tau_*/2 \ll 1$. This is a precise statement of the scale separation assumption which underlies Einstein's approach.

# 4 Ensemble averages and single realizations

In hydrodynamic dispersion, particles which begin at neighbouring points have similar histories in any single realization. Marbled endpapers in old books were produced by floating coloured inks on water, stirring the surface, and then capturing the swirls by carefully lowering a sheet of paper onto the inky film [3]. This technique, probably originating in Persia in the 1400s, presses hydrodynamic correlations into the service of art. Fortunately for printers, and distressingly for statisticians, a single realization does not resemble the blurry diffusion equation.

## 4.1 Eddy diffusion of a front

Figure 5 shows a single realization of the evolution of a "front" under the RW advection process. The front is the sharp border which separates white from dark; initially this line coincides with the $y$-axis. We suppose that the concentration is $c = -1$ for $x < 0$ and $c = +1$ for $x > 0$. Successive pulses of the renovating wave produce an increasingly folded front and the $c = -1$ fluid invades the region $x > 0$ in long thin tendrils. The central question is:

*How well is the process in figure 5 described by the diffusion equation?*

We know that given many realizations of this process, the long-time ensemble average of these realizations will follow the diffusion equation $\langle c\rangle_t = D\langle c\rangle_{xx}$, with the initial conditions $c(x, 0) = \pm 1$. The solution of this problem is

$$\langle c\rangle = \operatorname{erf}\eta\,, \qquad \text{where} \qquad \eta = \frac{x}{2\sqrt{Dt}}\,. \tag{25}$$

Figure 6 shows this smooth erf solution which, of course, looks nothing like figure 5. If the dark fluid in figure 5 contained radioactive contaminant, and we wanted to estimate the *maximum* exposure of at some value of $x > 0$, then the erf solution in (25) is not useful.

On the other hand, diffusivities are useful if we want to know how many particles are at such-and-such a distance from their initial location. Thus, figure 7 shows a histogram of the positions of 10,000 particles which all start on the line $x = 0$ (the initial front). The Gaussian curve in figure 7 is the corresponding prediction for the PDF of positions which is obtained by solving (19) with the initial condition $\langle c\rangle = \delta(x)$:

$$c(x, t) = \frac{1}{\sqrt{4\pi Dt}}\exp\left[-\frac{x^2}{4Dt}\right], \qquad D = \frac{\tau_*}{8}. \tag{26}$$

The histograms converge slowly to this Gaussian prediction. This asymptotic success shows that the diffusion equations correctly predicts the dispersion of particles when $t \gg \tau_*$.

An amusing aspect of the simple problem in figure 5 is that we can easily calculate the RMS fluctuations of $c$ around the ensemble average concentration in (25). Because $c = \pm 1$ we have $\langle c^2\rangle = 1$. Therefore, defining the fluctuation as

$$c' = c - \langle c\rangle\,, \tag{27}$$

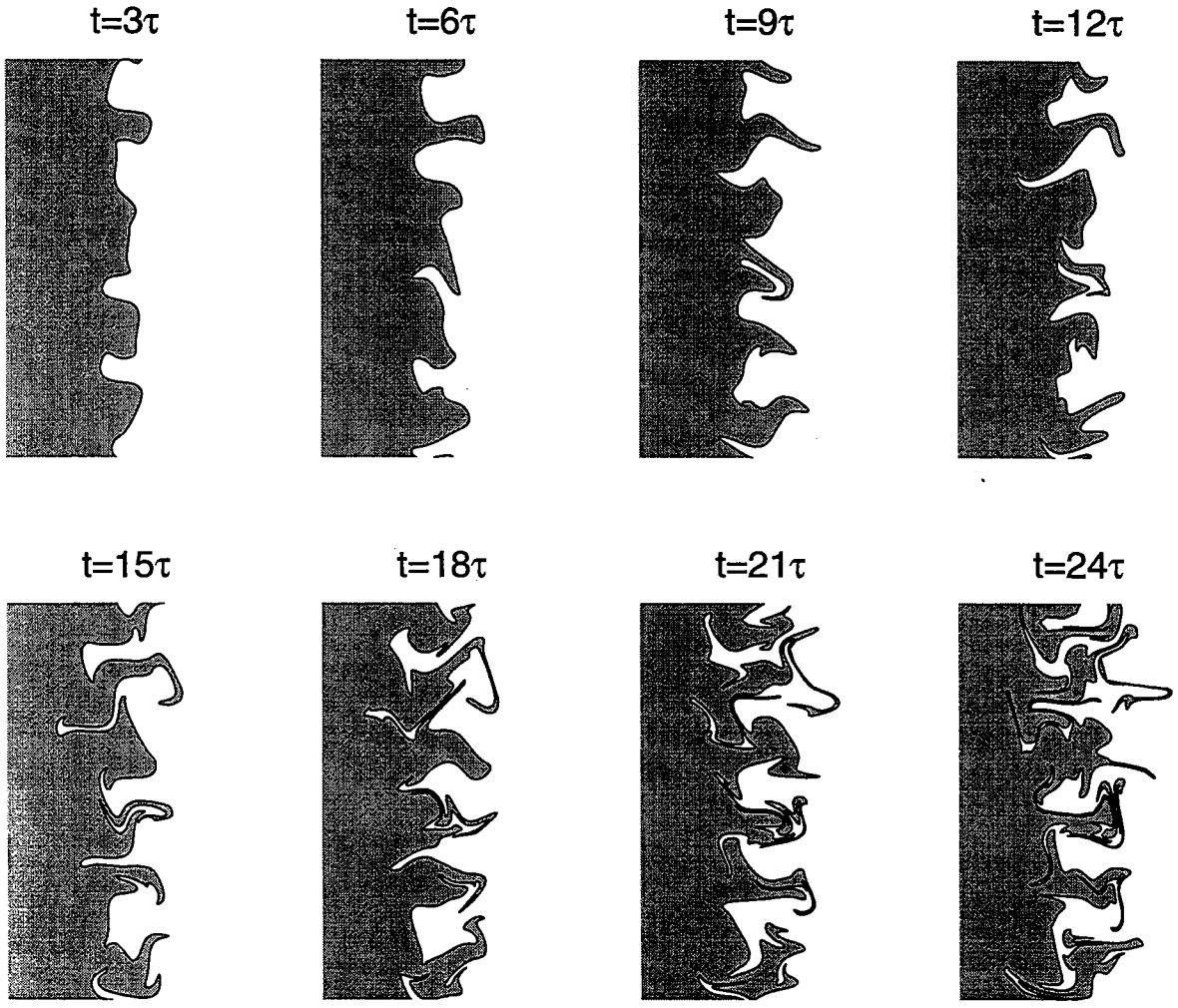


Figure 5: Evolution of a front under the advection by the RW model. The front initially coincides with the $y$-axis.

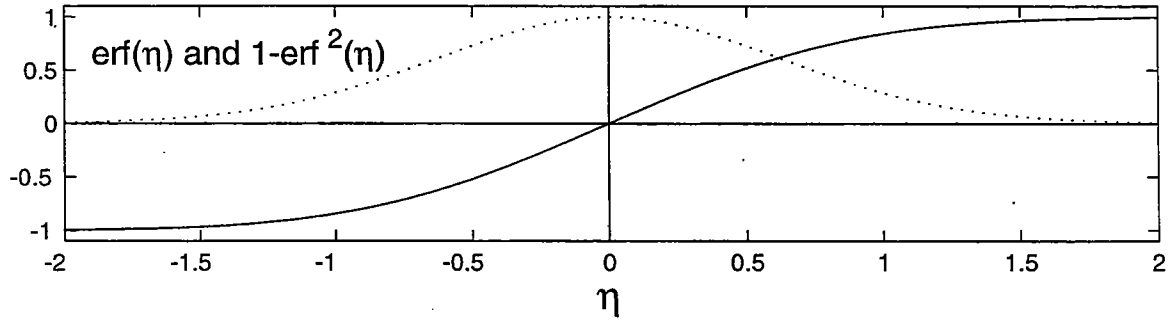


Figure 6: Evolution of the ensenble-averaged concentration $c$ and its variance during the evolution of the front underthe RRW model. Note how most of the variance is localised around $x = 0$.

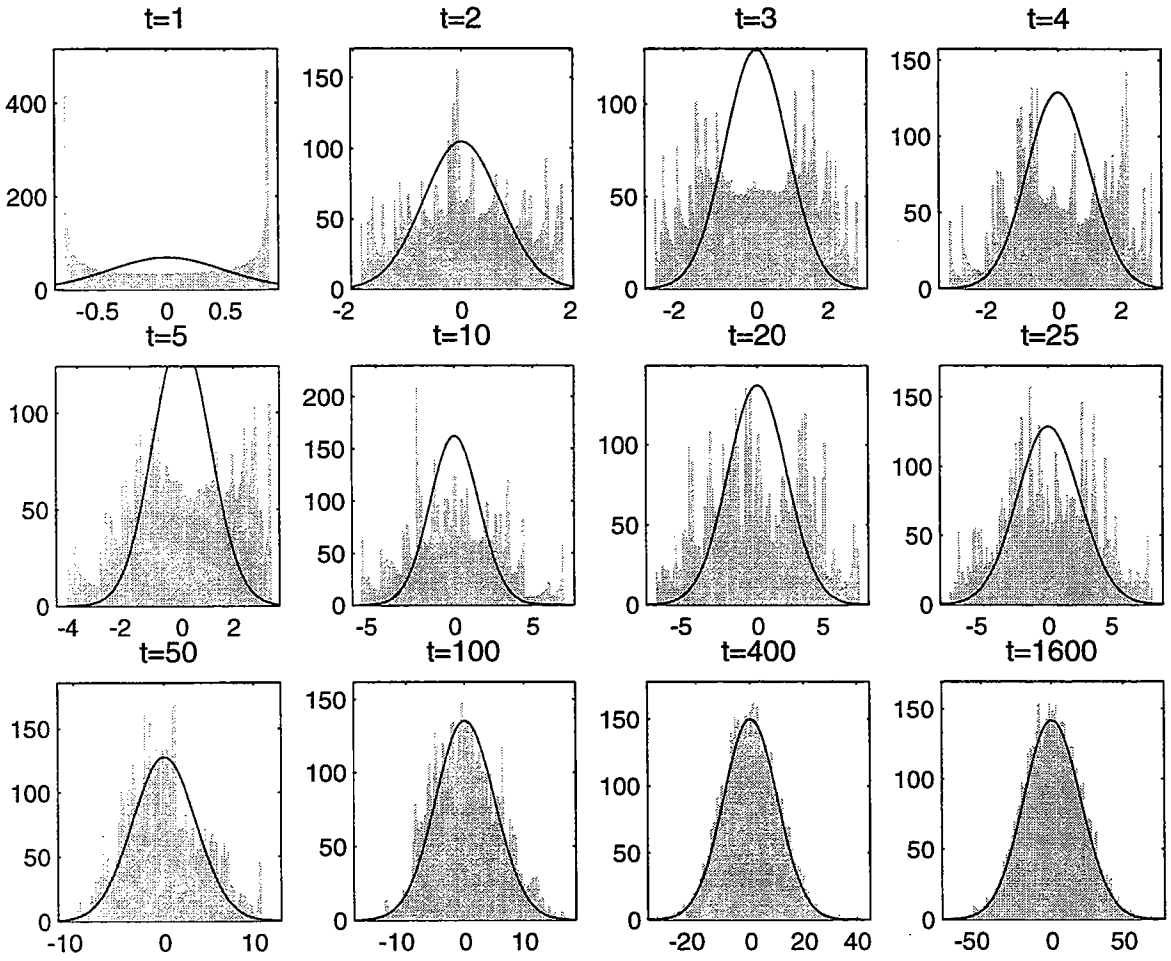


Figure 7: At $t = 0$ the front in figure 5 is tagged by placing 10000 particles along the $y$-axis. The historgram above shows the subsequent $x$-locations of these marker particles as the front is distorted by the RW model with $\tau_* = 1$. The Gaussian curve is given by (26).

we have

$$\langle c'^2 \rangle = \langle c^2 \rangle - \langle c \rangle^2 = 1 - \mathrm{erf}^2(\eta)\,. \tag{28}$$

The variance $\langle c'^2 \rangle$ is also indicated in figure 6.

## 4.2 Coarse grained averages and spatial filters

The process in figure 5 is translationally invariant in the $y$-direction and so using only a single realization we can calculate a spatially averaged concentration

$$\bar{c}(x,t) \equiv \lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} c(x,y,t)\mathrm{d}y\,. \tag{29}$$

The evolution of $\bar{c}$ will be asymptotically described by the diffusion equation.

In a general case, in which there is no statistical symmetry along a particular direction, one can take a single realization and define a coarse-grained or low-pass filtered concentration by:

$$\hat{c}(\mathbf{x},t) \equiv \int K(\mathbf{x}-\mathbf{x}')c(\mathbf{x}',t)\,\mathrm{d}^2\mathbf{x}'\,; \qquad K(|\mathbf{x}|)\text{is a filter.} \tag{30}$$

The hope is that scale separation between the width of the erf and the swirls will ensure that $\hat{c} \approx \langle c \rangle$. Thus the kernel of the filter, $K$ in (30), might be a Gaussian with a width which is at once much smaller than the thickness of the erf transition zone and much greater than an individual swirl in figure 5.

Scale separation is essential here because the filtering operation defined by the convolution in (30) is not strictly an "average". Some of the properties we take for granted when we use averages are

$$\langle c' \rangle = 0\,, \qquad \langle\langle c \rangle\rangle = \langle c \rangle\,, \qquad \langle\langle a \rangle\langle b \rangle\rangle = \langle a \rangle\langle b \rangle\,. \tag{31}$$

For the ensemble average, as indicated in (31), everything works.

For a filter, such as $\hat{}$ in (30), we can define the fluctuation concentration $c''$ in analogy with (27):

$$c'' \equiv c - \hat{c}\,. \tag{32}$$

But then $\widehat{c''} \neq 0$ and none of the other desiderata in (31) follow. In other words, spatial filtering instead of the ensemble averaging introduces a host of extra assumptions which should be carefully assessed (but almost never are).

## 4.3 A digression: Brownian bugs

I have hinted darkly at problems associated with spatial filters. These issues are largely ignored by optimistic scientists. The hope is that scale separation justifies the application of diffusive closures to the coarse-grained version of a single realization. Perhaps a justification of this optimistic approach is that the alternative seems so repellent. Nonetheless, it is important to realize that interpreting coarse-grained distributions as ensemble averages involves a nontrivial assumption. The best way of exposing this assumption is to exhibit a problem in which spatial filters and ensemble averages are very different. Accordingly, as a model of biological processes, we consider random walkers which both die and reproduce. We refer to these biological walkers as Brownian bugs.

The model is formulated by first placing $N \gg 1$ Brownian bugs randomly in the unit square; the boundary conditions are periodic in both directions. Each cycle of the simulation begins with a random walk step in which bug $k$, located at $\mathbf{x}_k = (x_k, y_k)$, is displaced to a new position

$$(x_k', y_k') = \mathrm{mod}\,[(x_k, y_k) + (\delta x_k, \delta y_k); 1]\,. \tag{33}$$

In (33), $\delta x_k$ and $\delta y_k$ are Gaussian random variables and the "mod" is to enforce the periodic boundary conditions and keep each bug in the unit square. After this random walk step, the second part of the cycle is a "coin toss" which results in either death (heads) or division (tails). When a lucky bug divides, the offspring is placed at the same position as the parent. This cycle of random displacement and random birth/death is repeated many times in order to simulate many generations of reproduction, death and dispersion.

The simulation shown in figure 8 was implemented in MATLAB using these rules. The striking result is that the density of bugs spontaneously develops large-scale clumps and voids. Figure 8 seems to show an inverse cascade of patch sizes: patches emerge on small scales in panel (b) and then, after more cycles, panels (c) and (d) show that the patches have expanded in scale. To quantify this impression, we have computed one-dimensional concentration spectra which show that an increasingly red spectrum develops.

A seemingly innocuous ingredient of the brownian-bug model is that deaths can occur anywhere, but births are always adjacent to a living bug. This asymmetry between birth and death is crucial for

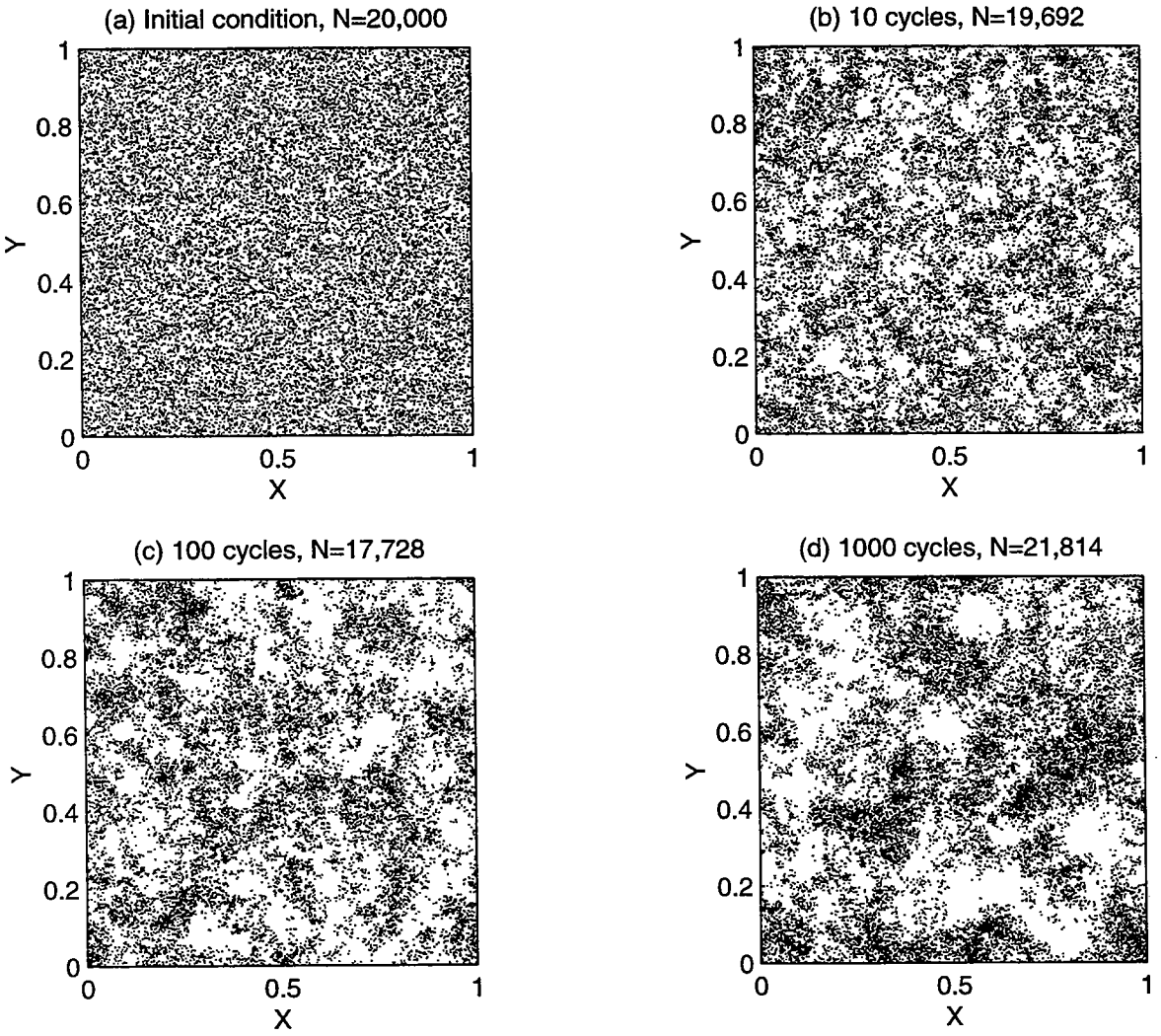


Figure 8: (a) The initial condition is $N = 20,000$ randomly located bugs in the unit square. Panels (b), (c) and (d) then show the development of patches after 10, 100 and 1000 cycles of random displacement followed by random birth/death. As the panel titles indicate, there are random fluctuations in the total size, $N$, of the population. The RMS step length of the underlying random walk is $\langle \delta x_k^2 \rangle^{1/2} = \langle \delta y_k^2 \rangle^{1/2} = 0.005$.

the spontaneous development of the voids and patches evident in figure 8: if one simulates birth by randomly placing the new bugs in the unit square then no patches form. This subtle point shows that making the births coincide with living bugs — surely a realistic feature of the model — introduces *pair* correlations. From another perspective, one can view the voids in figure 8 as the result of random extinctions which create voids. The step length of the random walk in figure 8 is such that diffusion is not strong enough to fill in the voids created by extinction.

The ensemble average of the Brownian bug process is described by

$$\langle c \rangle_t = D\nabla^2 \langle c \rangle + (\lambda - \mu)\langle c \rangle \,. \tag{34}$$

where $\lambda$ is the birthrate and $\mu$ the deathrate. However if the coin-toss is fair then births and deaths are equiprobable and consequently $\lambda = \mu$. In this case the solution of (34) which satisfies the initial

condition is

$$\lambda = \mu, \qquad \Longrightarrow \qquad \langle c \rangle = 1/N\,. \tag{35}$$

The uniform density above is the correct answer for the *ensemble* average concentration: the location of the voids and patches in figure 8 are accidently determined by the MATLAB random number generator. If we ensemble average many such patterns then the patches and voids must disappear because the process is spatially homogeneous.

On the other hand, the spatial average of a single realization, such as that in figure 8, will still show concentration patches[2]. Thus, in this Brownian bug example, $\hat{c} \neq \langle c \rangle$. Indeed, the patches are surely an important feature of the "real" answer. The correct but useless result in (35) exposes a failure of ensemble averaging. What do we make of this example? Are biological problems, with reproduction and death, so fundamentally different from the advection-diffusion of chemical tracers? I am not prepared to answer that question in these lectures and I leave further development of this example to the students.

# 5 Variance budgets

In this section we return to basics and present an alternative view of eddy-diffusivity. The following arguements emphasize the importance of the concentration variance equation.

## 5.1 The Reynolds' decomposition

Our point of departure is the advection-diffusion equation

$$c_t + \mathbf{u}\cdot\boldsymbol{\nabla} c = \kappa\nabla^2 c + s\,, \tag{36}$$

where $\kappa$ is the molecular diffusivity of $c$ and $\mathbf{u}$ is an incompressible ($\boldsymbol{\nabla}\cdot\mathbf{u} = 0$) velocity field. In (36) we have included a source term, $s(\mathbf{x}, t)$, which forces the system.

The velocity $\mathbf{u}$ in (36) is a single realization selected from an ensemble of velocity fields. Then we can introduce the "Reynolds' decomposition":

$$c = \langle c \rangle + c'\,, \tag{37}$$

where $\langle\,\rangle$ is the ensemble average and $c'$ is the fluctuation from $\langle c \rangle$ which arises in a single realization. Taking the ensemble average of (36) gives

$$\langle c \rangle_t + \langle \mathbf{u} \rangle\cdot\boldsymbol{\nabla}\langle c \rangle + \boldsymbol{\nabla}\cdot\langle \mathbf{u}'c' \rangle = \kappa\nabla^2\langle c \rangle + s\,. \tag{38}$$

(The source $s$ is taken to be deterministic, $\langle s \rangle = s$.)

Subtracting the ensemble average in (38) from (36) gives the fluctuation equation

$$c'_t + \langle \mathbf{u} \rangle\cdot\boldsymbol{\nabla} c' + \boldsymbol{\nabla}\cdot[\mathbf{u}'c' - \langle \mathbf{u}'c' \rangle] - \kappa\nabla^2 c' = -\mathbf{u}'\cdot\boldsymbol{\nabla}\langle c \rangle\,. \tag{39}$$

Equation (39) has been organized by taking the source term to the right hand side. Thus we see that advective distortion of the mean gradient, $\boldsymbol{\nabla}\langle c \rangle$, generates the fluctuation $c'$.

[2]If the width of the kernel, $K$ in (30), is larger than the dimension of the patches then filtering will remove the patches. However, since the patches expand in scale, eventually they will become so large that they survive the blurring power of the filter.

## 5.2 Consequences of linearity

If $c' = 0$ at $t = 0$ then, because (39) is linear, $c'$ and $\boldsymbol{\nabla}\langle c\rangle$ will be linearly related. It follows that the eddy flux $\langle \mathbf{u}'c'\rangle$ will also be linearly related to the mean gradient $\boldsymbol{\nabla}\langle c\rangle$. These simple observations, in alliance with the scale separation assumption, can be used to extract a surprising amount of information [2].

Because of the scale separation, it is plausible that this linear relation between eddy flux and mean gradient can be developed in a series of the form

$$\langle u_i'c'\rangle = -\mathcal{D}_{ij}^{(1)} * \langle c\rangle_{,j} - \mathcal{D}_{ijk}^{(2)} * \langle c\rangle_{,jk} + \cdots \tag{40}$$

The comma subscripts denote partial derivatives, $a_{,j} \equiv \partial a/\partial x_j$. We are also using the Einstein sum convention, where repeated indices are summed. The $*$ in (40) indicates that the product also involves convolutions in time, such as

$$\mathcal{D}_{ij}^{(1)} * \langle c\rangle_{,j} = \int_0^t \mathcal{D}_{ij}^{(1)}(t')\langle c\rangle_{,j}(t-t')\,\mathrm{d}t' \,. \tag{41}$$

If the mean field is varying slowly over an eddy decorrelation time then the convolution above approximates to

$$\langle u_i'c'\rangle \approx -\mathcal{D}_{ij}^{(1)} * \langle c\rangle_{,j} \approx -\int_0^\infty \mathcal{D}_{ij}^{(1)}(t')\,\mathrm{d}t'\,\langle c\rangle_{,j}(t) \,. \tag{42}$$

In the simplest cases[3]

$$\int_0^\infty \mathcal{D}_{ij}^{(1)}(t')\,\mathrm{d}t' = D_e\delta_{ij} \,, \tag{43}$$

where $D_e$ is the eddy diffusivity. Using (43) the flux gradient relation is

$$\langle \mathbf{u}'c'\rangle - \kappa\boldsymbol{\nabla}\langle c\rangle = -D\boldsymbol{\nabla}\langle c\rangle \,, \qquad D \equiv D_e + \kappa \,, \tag{44}$$

and the evolution of the average concentration is determined by

$$\langle c\rangle_t \approx D\nabla^2\langle c\rangle + s \,. \tag{45}$$

This is a general version of the specific diffusion equation derived in Section 3.3 for the renovating wave model.

## 5.3 The G · x-trick

The tensors $\mathcal{D}^{(n)}(t)$ are determined by the linear operator on the left-hand side of (39). Thus, these tensors depend on *(i)* the statistical properties of $\mathbf{u}'$; *(ii)* the mean advection $\langle \mathbf{u}\rangle$; *(iii)* the molecular diffusion $\kappa$. The essential point is that these tensors do not depend on $\langle c\rangle$. At least for the first term in the series, $\mathcal{D}_{ij}^{(1)}$, we can exemplify this by noting that there is a special solution of (36) in which $\langle \mathbf{u}\rangle = s = 0$ and concentration has the form

$$c = \mathbf{G}\cdot\mathbf{x} + c' \,. \tag{46}$$

[3] "Simple" means that the velocity ensemble is isotropic, homogeneous and reflexionally invariant. The last requirement means that the mirror image of a particular realization of $\mathbf{u}'$ is just as probable as $\mathbf{u}'$. If the ensemble is reflexionally invariant then $\mathcal{D}_{ij}^{(1)}$ is a symmetric tensor. This subtle point will be illustrated later in this lecture series.

In (46) the mean concentration is simply $\langle c \rangle = \mathbf{G} \cdot \mathbf{x}$ and the fluctuation $c'$ is determined from a reduced version of (39):

$$c'_t + \mathbf{u}' \cdot \boldsymbol{\nabla} c' - \kappa \nabla^2 c' = -\mathbf{G} \cdot \mathbf{u}' \,. \tag{47}$$

As emphasized above, the advection of the mean gradient appears as a source term for $c'$ on the right hand side of (47). Because (47) is linear, and $\mathbf{G}$ is constant, the solution $c'$ will be proportional to the large-scale gradient $\mathbf{G}$ and otherwise independent of $\mathbf{G}$.

This $\mathbf{G} \cdot \mathbf{x}$-trick enforces the platonic ideal of scale separation between the eddies and the mean field. If the concept of an eddy diffusivity is to have any validity, then it must work in the simplified context of (47). In fact, the $\mathbf{G} \cdot \mathbf{x}$-trick is used in doubly-periodic turbulence simulation to calculate eddy diffusivities. In that context, $\mathbf{u}' = (u, v)$ and $c'$ are efficiently represented by Fourier series. Then (47) is solved using a spectral code and the eddy flux is estimated by computing the integral

$$\langle \mathbf{u}' c' \rangle = A^{-1} \iint \mathbf{u}' c' \, \mathrm{d}x \, \mathrm{d}y \,, \tag{48}$$

over the computational domain. (In (48) $A$ is the total area of the domain so $\langle 1 \rangle = 1$). Notice that in (48) the ensemble average is identified with an integral over the domain. Later in these lectures we will use this same procedure to analytically calculate the eddy diffusivities of some spatially periodic velocity fields.

## 5.4 The concentration variance equation

An equation for the concentration variance,

$$\mathcal{Z} \equiv \frac{1}{2} \langle c'^2 \rangle \,, \tag{49}$$

is obtained by multiplying (39) by $c'$ and ensemble averaging. The result is

$$\mathcal{Z}_t + \langle \mathbf{u} \rangle \cdot \boldsymbol{\nabla} \mathcal{Z} + \boldsymbol{\nabla} \cdot \langle \frac{1}{2} \mathbf{u}' c'^2 \rangle - \kappa \nabla^2 \mathcal{Z} = -\kappa \langle \boldsymbol{\nabla} c' \cdot \boldsymbol{\nabla} c' \rangle - \langle \mathbf{u}' c' \rangle \cdot \boldsymbol{\nabla} \langle c \rangle \,. \tag{50}$$

The terms on the left-hand side of (50) can be interpreted as fluxes of $\mathcal{Z}$. The two terms on the right hand side of (50) are respectively dissipation of variance by molecular diffusion, $\kappa$, and a source of variance due to advective distortion of the mean gradient.

## 5.5 Heuristic closure arguments

In (50) there are three terms which we would like to relate to the mean quantities $\langle c \rangle$ and $\mathcal{Z}$. First, there is $-\langle \mathbf{u}' c' \rangle \cdot \boldsymbol{\nabla} \langle c \rangle = D_e \boldsymbol{\nabla} \langle c \rangle \cdot \boldsymbol{\nabla} \langle c \rangle$. The remaining two terms are $\langle \mathbf{u}' c'^2 / 2 \rangle$ and $\kappa \langle \boldsymbol{\nabla} c' \cdot \boldsymbol{\nabla} c' \rangle$.

The correlation $\langle \mathbf{u}' c'^2 / 2 \rangle$ in (50) is an eddy-flux of $c'^2$, just as $\langle \mathbf{u} c' \rangle$ is an eddy flux of $c'$. Thus, by analogy with (44), we can argue that

$$\frac{1}{2} \langle \mathbf{u}' c'^2 \rangle = -D_e \boldsymbol{\nabla} \mathcal{Z} \,. \tag{51}$$

This heuristic argument is discussed further in appendix B.

The final term in (50) is the dissipation of variance by molecular diffusivity, $\kappa \langle \boldsymbol{\nabla} c' \cdot \boldsymbol{\nabla} c' \rangle$. The simplest closure assumption we can make about this term is that

$$\kappa \langle \boldsymbol{\nabla} c' \cdot \boldsymbol{\nabla} c' \rangle \approx \beta \mathcal{Z} \,, \tag{52}$$

where $\beta$ has the dimensions of time. The closure above relies on dimensional analysis and the linearity of (36). However, in anticipation of a later discussion of the Batchelor spectrum, we now make some heuristic arguments in support of (52) which suggest that $\beta$ is independent of the molecular diffusivity.

Suppose that the mean field $\langle c \rangle$ has a length scale $L$ and that the velocity field $\mathbf{u}'$ has a length scale $L_{\mathbf{u}}$ (in the RW example $L_{\mathbf{u}} = k^{-1}$). The scale separation assumption is that

$$L \gg L_{\mathbf{u}} \,. \tag{53}$$

The inequality in (53) is exemplified in idealized case of (46) in which $L$ is infinite. If follows that advective distortion of $\boldsymbol{\nabla}\langle c \rangle$ generates $c'$ first on the scale $L_{\mathbf{u}}$. Then, following our arguments in lecture 1, the scale of $c'$ will be exponentially reduced, like $\exp(-\gamma t)$, where $\gamma$ is roughly proportional to the RMS strain of $\mathbf{u}'$. This exponential contraction continues until the *cascade* is halted by molecular diffusion at the scale

$$\ell \equiv \sqrt{\frac{\kappa}{\gamma}} \,. \tag{54}$$

Using arguments from lecture 1, we can estimate that the time taken for this arrest at $\ell$ is

$$t_\ell \approx \gamma^{-1} \ln \left( L_{\mathbf{u}} / \ell \right) \,. \tag{55}$$

Then the smallest length scale in the $c'$-field is $\ell$ and, plausibly, the gradient is $\boldsymbol{\nabla} c' \sim c'_{RMS}/\ell$ where $c'_{RMS} \equiv \sqrt{2\mathcal{Z}}$. We now have a simple estimate $\kappa \langle \boldsymbol{\nabla} c' \cdot \boldsymbol{\nabla} c' \rangle \sim \gamma \mathcal{Z}$. This rough argument leads to the closure in (52), with $\beta \propto \gamma$, and the caveat that $t > t_\ell$.

We can summarize the arguments above by rewriting the variance equation (50) as

$$\mathcal{Z}_t + \langle \mathbf{u} \rangle \cdot \boldsymbol{\nabla} \mathcal{Z} - D \nabla^2 \mathcal{Z} = D_e \boldsymbol{\nabla} \langle c \rangle \cdot \boldsymbol{\nabla} \langle c \rangle - \beta \mathcal{Z} \,, \qquad (\text{if } t \geq t_\ell) \,. \tag{56}$$

The most dubious approximation is probably (52). To conclude this discussion we will interpret the variance equation in two specific examples.

## 5.6 Example 1: the dispersing front

First consider the dispersing front in figure 5. In this example $s = \kappa = \langle \mathbf{u} \rangle = 0$ and we have already know from (28) that

$$\mathcal{Z} = \frac{1}{2} \left[ 1 - \mathrm{erf}^2 \left( \eta \right) \right] \,, \qquad \eta = \frac{x}{2\sqrt{Dt}} \,. \tag{57}$$

On the other hand, since $\kappa = 0$, it follows that $D = D_e$ and $\beta = 0$. With these simplifications the variance equation (50) reduces to

$$\mathcal{Z}_t - D \mathcal{Z}_{xx} = D \boldsymbol{\nabla} \langle c \rangle \cdot \boldsymbol{\nabla} \langle c \rangle \,, \tag{58}$$

where $\langle c \rangle$ is the erf-solution in (25). As a consistency check, one can show that (57) is the solution of the variance equation in (58).

This example shows that the destruction of variance by molecular diffusivity is not required in order to prevent an accumulation of variance: the source on the right-hand side of (58) is balanced by eddy diffusion.

## 5.7 Example 2: a large-scale source

In this second example the tracer is injected by a source $s = \cos qx$ in (36). We also take $\langle \mathbf{u} \rangle = 0$ so that the mean concentration field is obtained by solving

$$\langle c \rangle_t - D \nabla^2 \langle c \rangle = \cos qx \,, \quad \Longrightarrow \quad \langle c \rangle = \frac{1}{Dq^2} \left[ 1 - e^{-Dq^2 t} \right] \cos qx \,. \tag{59}$$

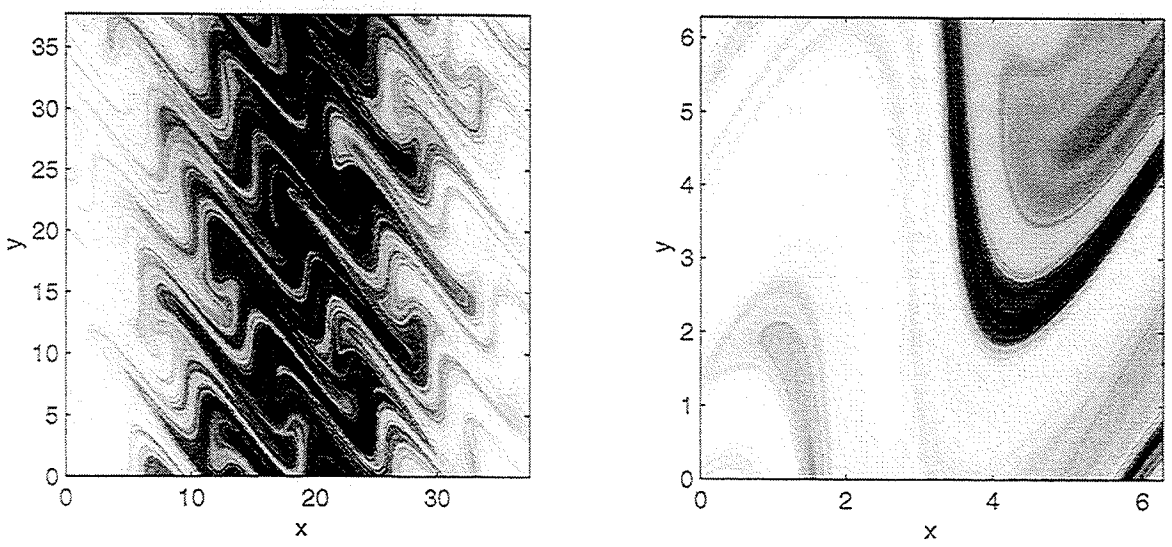


Figure 9: A simulation of the source problem, with $s = \cos(x/6)$, using the RW model to generate $\mathbf{u}$ in (36). There is no molecular diffusivity ($\kappa = 0$). The left-hand panel shows the whole domain (the length of the side is $12\pi$) while the right hand panel shows a smaller subdomain (the length of the side is $2\pi$). The concentration fields were generated by 10 pulses of the renovating wave using $\tau_* = 3$ (that is, $t = 30$).

(To apply the diffusion equation the scale of the source, $q^{-1}$, must be much larger than the scale of the velocity field.) A steady mean concentration pattern is established when $Dq^2t \gg 1$.

The concentration variance is determined by solving the variance equation (56)

$$\mathcal{Z}_t - D\nabla^2\mathcal{Z} = \frac{1}{2}\frac{D_e}{D^2q^2}\left[1 - e^{-Dq^2t}\right]^2 (1 - \cos 2qx) - \kappa\langle\boldsymbol{\nabla}c'\cdot\boldsymbol{\nabla}c'\rangle\,. \tag{60}$$

In (60), the solution in (59) has been used to evaluate the source term on the right hand side and we have left the diffusive sink in its exact form.

It is clear from (60) that the molecular diffusion, $\kappa$, plays an important role. If $\kappa = 0$ then the long time solution of (60) has a component which eventually grows linearly with time:

$$\kappa = 0\,, \qquad \Longrightarrow \qquad \mathcal{Z} \propto t/2Dq^2\,. \tag{61}$$

Thus, without molecular diffusion, there is "runaway variance". Ultimately, in a single realization, the mean field in (59) will be buried under enormous fluctuations.

To give an intuitive derivation of (61) we argue that with $\kappa = 0$ the concentration on each fluid element is determined by solving the Lagrangian equation

$$\frac{Dc}{Dt} = \cos qx(t)\,, \tag{62}$$

where $x(t)$ is the randomly changing $x$-position of the particle. Thus, the concentration on each particle is undergoing a random walk along the $c$-axis, which is induced by the random motion of the particle through the $\cos qx$ source function. The decorrelation time of this walk is the time it takes a particle to diffuse through a distance of order $q^{-1}$, which is $1/Dq^2$. Thus, in a time $t$, there are roughly $N(t) \sim Dq^2t$ independent steps along the $c$-axis. But because the source acts coherently for a time $1/Dq^2$ with a strength of order unity, the step length of this random walk is roughly $\Delta c \sim 1/Dq^2$. Thus, the mean square displacement of $c$ is:

$$\langle c'^2\rangle \sim (\Delta c)^2 N(t) \sim \frac{t}{Dq^2}\,, \tag{63}$$

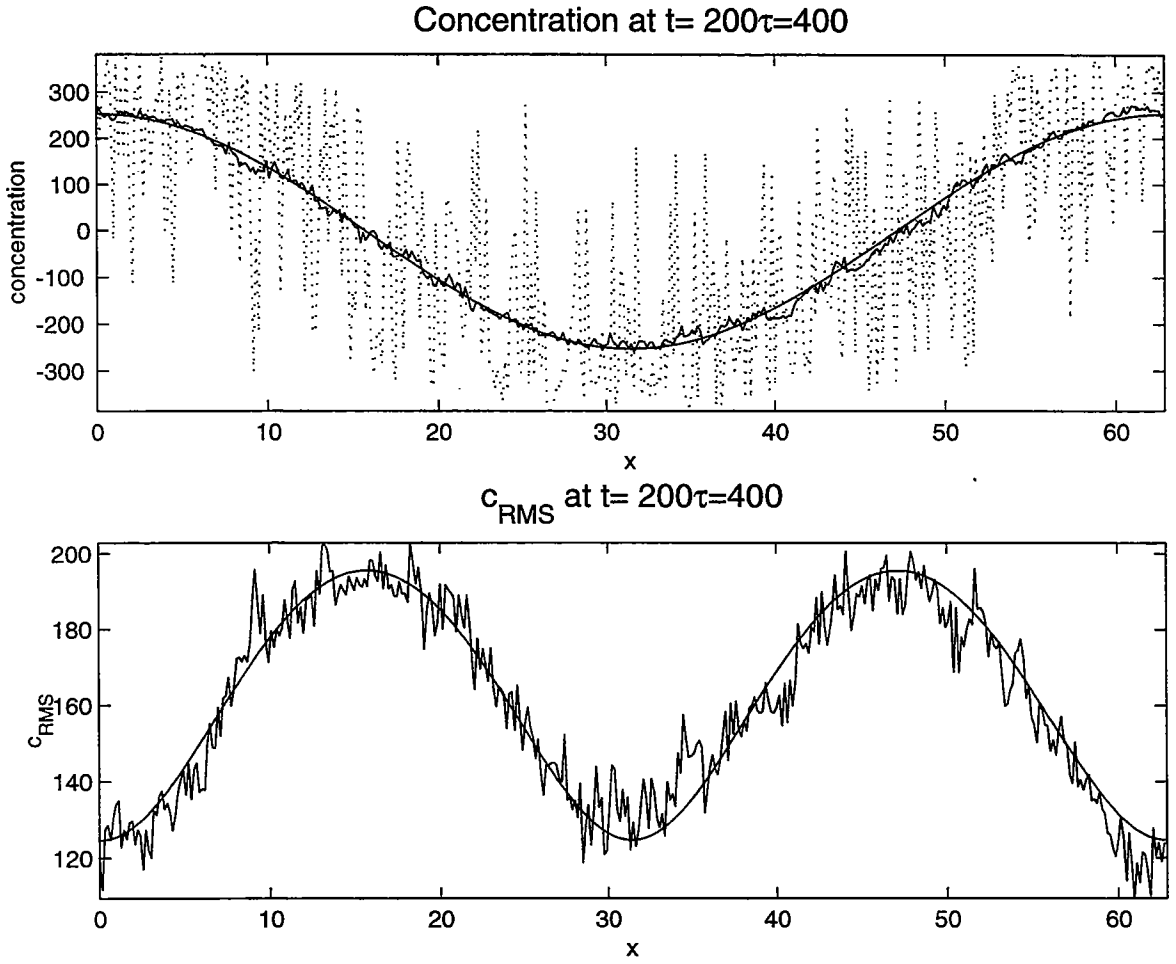


Figure 10: This figure compares analytic results with a numerical solution of (62), taking $q = 1/10$, and using the RW model to generate $\mathbf{u}$. The persistence parameter is $\tau_* = 2$ and the results are at $t = 400$ (that is, 200 renovation cycles). The concentration $c(\mathbf{x}, 400)$ is calculated on a $400 \times 400$ grid using the method in appendix B. In the top panel there are three curves: the concentration as a function of $0 < x < 20\pi$ along the line $y = 0$ (the jagged dotted curve); the $y$-averaged concentration defined in (64); the analytic result in (59) (the smooth sinusoid). The bottom panel compares the $c_{RMS} = \sqrt{2\mathcal{Z}}$ obtained by solving (60) analytically with $c_{RMS}$ estimated using (64).

which is the final result in (61).

It is interesting to compare the analytic results in (59) and (60) with a numerical solution of (62). Thus we must compute the spatial averages

$$\bar{c}(x,t) \equiv \frac{1}{L}\int_0^L c(x,y,t)\,\mathrm{d}y, \qquad c^2_{RMS}(x,t) \equiv \frac{1}{L}\int_0^L [c(x,y,t)-\bar{c}]^2\,\mathrm{d}y\,, \tag{64}$$

using the numerical solution, and compare these with the analytic results for $\langle c\rangle$ and $\mathcal{Z} = c^2_{RMS}/2$. The best way to make this comparison is to obtain $c(x,y,t)$ on a regular grid in the $(x,y)$-plane. As a bonus, one can then also use contouring routines to make pretty pictures of the concentration field (see figure 9).

The concentration field is calculated on a regular grid using the procedure described in Appendix C

(essentially the method of characteristics). Figure 10 shows good agreement between this simulation and analytic results. Notice that in figure 10 the variance $\mathcal{Z}$ peaks where $\boldsymbol{\nabla}\langle c\rangle$ is greatest. This illustrates that concentration fluctuations are produced by advective distortion of the mean gradient: where the mean gradient is large there is lots of variance. But $\mathcal{Z} \neq 0$ even where $\boldsymbol{\nabla}\langle c\rangle = 0$ (for example, at $x = 0$ and $x = 10\pi$ in figure 10). Thus, where the source term on the right hand side of (60) vanishes, the diffusive term $D\nabla^2\mathcal{Z}$ supplies variance.

### 5.8 Cautionary remarks

In the both examples above there is no molecular diffusion ($\kappa = 0$) and consequently there is no destruction of variance by the term $\kappa\langle\boldsymbol{\nabla}c'\cdot\boldsymbol{\nabla}c'\rangle$ in (50). As a project for a student, include molecular diffusion in the RW model (perhaps by pulsing diffusion in alternation with advection) and assess the efficacy of this process. In particular, can the closure in (52) be justified?

## References


[1] S. Childress and A. D. Gilbert. *Stretch, Twist, Fold: The Fast Dynamo.* Springer, Berlin, 1995.

[2] H.K.Moffatt. Transport effects associated with turbulence with particular attention to the influence of helicity. *Rep. Prog. Phys.*, 46:621–664, 1983.

[3] W.A. Medeiros. *Marbling Techniques.* Watson-Guptill, New York, 1994.

[4] G.I. Taylor. Eddy motion in the atmosphere. *Phil. Trans. Roy. Soc. London*, 215:1–16, 1915.


## A Calculation of the RW Green's function

In this appendix we present an alternative calculation of the RW ensemble averaged Green's function, $g(r)$, in (17). The unaveraged Green's function, $G(\mathbf{x}, \mathbf{x}_0, t)$, is the solution of (10). Because the process is spatially homogeneous it is harmless to take $\mathbf{x}_0 = 0$ so that

$$G(\mathbf{x}, 0, \tau_*) = \delta\left[x - \tau_* s\sin\varphi\right]\delta\left[y + \tau_* c\sin\varphi\right], \tag{65}$$

where $(s, c) \equiv (\sin\theta, \cos\theta)$. The ensemble average of (65) is computed by integration over $\varphi$ and $\theta$, as in (4). It is very pleasant that there are two integrals and two $\delta$-functions. Thus, we first do the $\varphi$-integral by noting that $\delta\left[x - \tau_* s\sin\varphi\right]$ is nonzero at the two values of $\varphi$ where $\sin\varphi = x/\tau_* s$, and at those positions:

$$\frac{\mathrm{d}}{\mathrm{d}\varphi}\left[x - \tau_* s\sin\varphi\right] = \pm\sqrt{\tau_*^2 s^2 - x^2}\,. \tag{66}$$

Using the standard rule for changing variables in a $\delta$-function, we find that the average of (65) over $\varphi$ alone is

$$\langle G\rangle_\varphi = \frac{1}{\pi}\frac{\delta(y + \cot\theta x)}{\sqrt{\tau_*^2\sin^2\theta - x^2}}\,. \tag{67}$$

The second integral over $\theta$ is performed by noting that $\delta(y + \cot\theta x)$ is nonzero at the two values of $\theta$ where $\cot\theta = -y/x$, and at those positions

$$\sin^2\theta = \frac{x^2}{x^2 + y^2}\,, \qquad \frac{\mathrm{d}}{\mathrm{d}\theta}\left[y + x\cot\theta\right] = -\frac{x^2 + y^2}{x}\,. \tag{68}$$

After changing variables in the $\delta$-function we recover $g(r)$ in (17).

# B Eddy diffusion of variance

Ignoring small molecular diffusion ($\kappa = 0$), if $c$ satisfies the advection equation then any function of $c$ satisfies the same equation. That is to say

$$\frac{Dc}{Dt} = 0, \qquad \Longrightarrow \qquad \frac{Df}{Dt} = 0\,, \tag{69}$$

where $f(c) = c^2$, or $\exp(c)$, etcetera. Taking an ensemble average, and making the same arguments for $f(c)$ as for $c$, we have that

$$\langle f \rangle_t = D\nabla^2 \langle f \rangle\,. \tag{70}$$

In the particular case $f = c^2/2$, $\langle f \rangle = \langle c \rangle^2/2 + \mathcal{Z}$ and (70) reduces to

$$\mathcal{Z}_t = D\nabla^2 \mathcal{Z} + D\boldsymbol{\nabla}\langle c \rangle \cdot \boldsymbol{\nabla}\langle c \rangle\,. \tag{71}$$

Matching the terms in (71) with those in (50) we conclude that $\langle \mathbf{u}' c'^2/2 \rangle = -D\boldsymbol{\nabla}\mathcal{Z}$.

# C Numerical simulation of the RW process

Drawing figures 9 and 10 requires that we obtain the solution of (62) on a regular grid in the $(x, y)$-plane. This is an opportunity to use the method of characteristics and learn some MATLAB programming techniques.

Equation (3) shows how the movement of a particle in the RW velocity field is equivalent to a random map which determines the position at $(n+1)\tau_*$ in terms of the previous position at $n\tau_*$. If this particle carries a concentration, $c(\mathbf{x}, t)$, which changes because of the $\cos qx$ source in (62), then the concentration changes can also be calculated and expressed as a map in discrete time.

Thus, suppose that the concentration on a particle at time $t = n\tau_*$ is $c_n$. Then the change in concentration during $n\tau_* < t < (n+1)\tau_*$ is obtained by integrating

$$\frac{Dc}{Dt} = \cos\left[qx_n + qu_n(t - n\tau_*)\right]\,, \tag{72}$$

where the constant $x$-velocity of the particle is $u_n = s_n \sin(c_n x_n + s_n y_n + \varphi_n)$, with $(s_n, c_n) \equiv (\sin\theta_n, \cos\theta_n)$. The integral of (72) can be written as

$$c_{n+1} = c_n + \frac{\sin(qx_{n+1}) - \sin(qx_n)}{qu_n}\,. \tag{73}$$

With equations (73) and (3) we can advance forward in time and so determine the concentration on a particle at $t = n\tau_*$.

However we need to determine the concentration at $t = n\tau_*$ at a specified grid point $\mathbf{x}$. The trick is illustrated in the Matlab program below.

```
%% Solution of
%%
%%   Dc/Dt=cos(qx);
%%
%%  cos(q x) is a large-scale source and u is the RW velocity.
%%  The RW streamfunction is psi=cos[cos(theta) x+ sin(theta)y + phi]
clc
```

```
N=400;              %% Use an N*N grid in the plotting window
q=1/6;              %% The wavenumber of the cos q x source
LL=2*pi/q;          %% LL is the domain size
npulse=10           %% The number of renovation cycles
tau=3;              %% The pulse duration of the wave

%% Lwin is the side of the square plotting window.
%% Set Lwin=LL to see the big picture. To see small scale details,
%% try Lwin = 2*pi. We draw two subplots with different Lwin's
nloop=0;
 for Lwin=[LL 2*pi]
   nloop=nloop+1
       x=linspace(0,Lwin,N);   %% x is the coordinate in the plotting window.
       h=x(2);                 %% The grid spacing in the plotting window

         for j=1:N
          jj=[((j-1)*N+1):(j*N)];
          pos(jj,1)=zeros(N,1)+(j-1)*h;
          pos(jj,2)=x';
         end

       conc=zeros(N*N,1);

  %% The position of the N^2 particles are stored in pos with
  %% N^2 rows and 2 columns. Each vertical segment of
  %% length N in pos contains particles with the same initial x-position.
  %% the column vector conc contains the concentration on the
  %% N*N particles in pos. Initially, conc=0 at the N*N
  %%  grid points. Then we integrate
  %% backwards in time to find the concentration change.

     for k=1:1:npulse
       theta=rand*2*pi;
       wavevec=[cos(theta),sin(theta)]';
           phase=rand*2*pi;
       vel=sin(pos*wavevec+phase)*[wavevec(2),-wavevec(1)];
       conc=conc-sin(q*pos(:,1))./(q*vel(:,1));
       pos=pos+tau*vel;
       conc=conc +sin(q*pos(:,1))./(q*vel(:,1));
     end

%% Emerging from this loop, we have the the new positions
%% and the new concentration

  conc=reshape(conc,N,N); %% conc is reshaped into an N*N matrix
  hh=subplot(1,2,nloop)
  colormap('gray')
  imagesc(x,x,conc)
  axis equal
  xlabel('x')
  ylabel('y')
```

```
  axis([0 Lwin 0 Lwin])
  set(hh,'ydir','norm')

end
```

# Lecture 3: Stretching

## 1 Line stretching

In the previous lecture we emphasized that the destruction of tracer variance by molecular diffusivity relies on the increase of $\nabla c$ by stirring. Thus the term $\kappa\langle\nabla c'\cdot\nabla c'\rangle$ in the variance budget eventually becomes important, even though the molecular diffusivity $\kappa$ is very small. One goal of this lecture is to understand in more detail how tracer gradients in a moving fluid are amplified by simple velocity fields. We will assume that $\kappa = 0$ so that there is stirring without mixing. This is a good approximation provided that the smallest scale in the tracer field is much greater than the length $\ell = \sqrt{\kappa/\alpha}$ that we identified in lecture 1.

Gradient amplification is closely related to the stretching of material lines, a subject that was opened by Batchelor in 1952. A material line is a curve that consists always of the same fluid particles. Batchelor's main conclusion is that there is a timescale governing the ultimate growth of an infinitesimal line element, but no length scale other than that of the element itself. These dimensional considerations force the conclusion that the element grows exponentially,

$$\ell = \ell_0 \mathrm{e}^{\gamma t}\,, \tag{1}$$

where $\gamma$ is a constant with dimensions of inverse time, related to the timescale that Batchelor had in mind.

Just as some close particle pairs separate exponentially, other pairs starting at distant points are brought close together. This might seem paradoxical until one recalls the folded tracer patterns evident in Welander's 1955 experiments (see the final figures in lecture 1). If two closely approaching particles are carrying different values of $c$ then the gradient $\nabla c$ will be amplified. Thus, as a corollary of (1) we expect that $|\nabla c| \sim |\nabla c_0| \exp(\gamma t)$. It is through this exponential amplification of the concentration gradients that the small molecular diffusivity $\kappa$ is able eventually to destroy tracer variance.

### 1.1 Material line elements and tracer gradients

Using a geometric argument (see figure 1) we can give a proof-by-intimidation that a material line element, $\boldsymbol{\xi}(\boldsymbol{x}, t)$, attached to a fluid element evolves according to

$$\frac{D\boldsymbol{\xi}}{Dt} = (\boldsymbol{\xi}\cdot\nabla)\boldsymbol{u}\,. \tag{2}$$

Here the "convective derivative" is $D/Dt = \partial/\partial t + \boldsymbol{u}\cdot\nabla$. The field of line elements can be visualized a collection of tiny straight arrows attached to each moving particle of fluid. Then (2) describes the evolution of this collection of arrows. Notice that (2) refers to an *infinitesimal* line element $\boldsymbol{\xi}$. If the length of a material line is comparable to the scale of $\boldsymbol{u}$ there is no longer a simple relation between the stretching of the material line and *local* properties of $\boldsymbol{u}$, such as $\nabla\boldsymbol{u}$.

Taking the gradient of the tracer equation

$$\frac{Dc}{Dt} = 0\,, \tag{3}$$

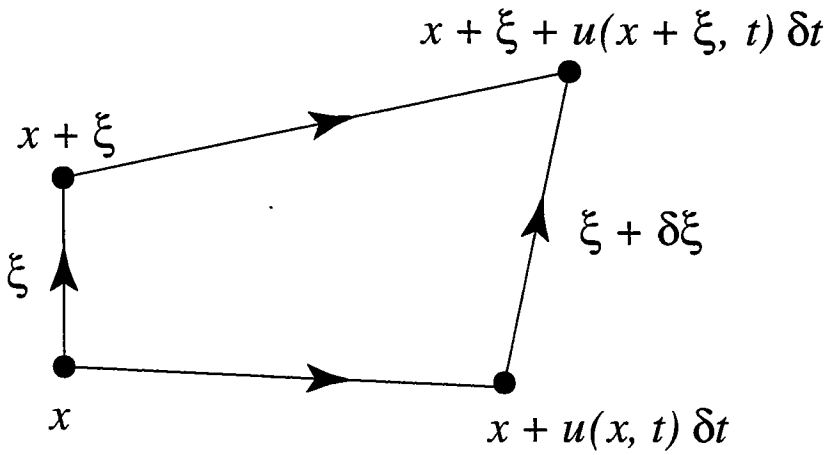


Figure 1: The line element $\boldsymbol{\xi}$ is short enough to remain straight and to experience a strain that is uniform over its length during the time $\delta t$. Proof by intimidation of (2) : $\delta\boldsymbol{\xi} = [\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{\xi},t) - \boldsymbol{u}(\boldsymbol{x},t)]\delta t$, and take $(\delta t, \boldsymbol{\xi}, \delta\boldsymbol{\xi}) \to 0$.

gives

$$\frac{D\nabla c}{Dt} = -(\nabla c\cdot\nabla)\boldsymbol{u}\,. \tag{4}$$

Despite the difference in the sign of the right hand sides of (2) and (4) there is a close connection between the solutions of the two equations.

To emphasize the connection between $\nabla c$ and $\boldsymbol{\xi}$, we mention the conservation law

$$\frac{D}{Dt}(\nabla c\cdot\boldsymbol{\xi}) = 0\,. \tag{5}$$

(Meteorologists and oceanographers might recognize (5) as a relative of potential vorticity conservation.) Later in this lecture (5) is used to deduce $\nabla c$ from $\boldsymbol{\xi}$.

The easy way to prove (5) is to consider a pair of particles separated by a small displacement $\boldsymbol{\xi}$. If the concentration carried by the first particle is $c_1$, and that of the second particle is $c_2 = c_1 + \mathrm{d}c$, then $\mathrm{d}c = \boldsymbol{\xi}\cdot\nabla c$. Thus (5) is equivalent to the "obvious" fact that $\mathrm{d}c$ is conserved as the two particles move.

The difficult way to prove (5) is to take the dot product of $\nabla c$ with (2) and add this to the dot product of $\boldsymbol{\xi}$ with (4). Performing some nonobvious algebra, perhaps with Mathematica or Maple, one can eventually simplify the mess to (5). Suffering through this tedious exercise will convince the student that the earlier, easy proof is worthy of serious attention.

## 1.2 Eulerian versus Lagrangian: the golden rule

Particle trajectories, $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{x}_0, t)$, are determined by solving the differential equations

$$\frac{D\boldsymbol{x}}{Dt} = \boldsymbol{u}(\boldsymbol{x},t)\,, \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0\,. \tag{6}$$

The solution of the differential equation above defines the particle position, $\boldsymbol{x}$, as a function of the two independent variables, $\boldsymbol{x}_0$ and $t$. Using this time-dependent mapping between $\boldsymbol{x}$ and $\boldsymbol{x}_0$, we can take a problem posed in terms of $\boldsymbol{x}$ and $t$ (the Eulerian formulation) and change variables to obtain an equivalent formulation in terms of $\boldsymbol{x}_0$ and $t$ (the Lagrangian formulation).

In the Eulerian view, the independent variables are $\boldsymbol{x} = (x, y, z)$ and $t$. The convective derivative,

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}\,, \tag{7}$$

is a differential operator involving all of the independent variables.

In the Lagrangian view, the independent variables are $\boldsymbol{x}_0$ and $t'$ and $\boldsymbol{x}(\boldsymbol{x}_0, t')$ is a dependent variable. As an accounting device, the time variable is decorated with a prime to emphasize that a $t'$-derivative means that the independent variables are $\boldsymbol{x}_0$. To move between the Eulerian and Lagrangian representations notice that

$$\frac{\partial t}{\partial t'} = 1\,, \quad \text{and} \quad \frac{\partial}{\partial t'}(x, y, z) = (u, v, w)\,. \tag{8}$$

The second equation above is the definition of velocity, $\boldsymbol{u} = (u, v, w)$.

Using (8), the rule for converting partial derivatives is

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + \frac{\partial x}{\partial t'}\frac{\partial}{\partial x} + \frac{\partial y}{\partial t'}\frac{\partial}{\partial y} + \frac{\partial z}{\partial t'}\frac{\partial}{\partial z} = \frac{D}{Dt}\,. \tag{9}$$

Equation (9) is the golden rule that enables us to interpret expressions such as

$$\frac{D}{Dt}(\text{unknown}) = \text{RHS} \tag{10}$$

in either Eulerian or Lagrangian terms. Using the golden rule we can dispense with the prime that decorates the Lagrangian time variable.

In the Eulerian interpretation we must express the RHS in (10) as a function of $x$, $y$, $z$ and $t$ and use the Eulerian definition of the convective derivative in (7). Then (10) is a *partial* differential equation for the unknown.

In the Lagrangian interpretation $D/Dt$ is the same as a simple time derivative and we must express the RHS of (10) as a function of $x_0$, $y_0$, $z_0$ and $t$. Then (10) is a *ordinary* differential equation for the unknown.

As an illustration of the transformation between Eulerian and Lagrangian variables, consider the steady, unidirectional velocity field $\boldsymbol{u} = [u(y), 0]$. The solution of (6) is

$$x = x_0 + u(y)t\,, \qquad y = y_0\,. \tag{11}$$

In this example it is a simple matter to express $(x, y)$ in terms of $(x_0, y_0)$ and vice versa.

The line-stretching equation, (2), has the same form as (10). For the same unidirectional velocity field, using components, $\boldsymbol{\xi} = (\xi, \eta)$, we have in Lagrangian variables

$$\frac{D\xi}{Dt} = \eta u'(y_0)\,, \qquad \frac{D\eta}{Dt} = 0\,. \tag{12}$$

(We have used the golden rule (9) above.) Equation (12) is an ordinary differential equation and the solution is

$$\xi = \xi_0(x_0, y_0) + t\eta_0(x_0, y_0)u'(y_0)\,, \qquad \eta = \eta(x_0, y_0)\,. \tag{13}$$

Using (11), we can write (13) in terms of Eulerian variables as

$$\xi = \xi_0[x - u(y)t, y] + t\eta_0[x - u(y)t, y]u'(y)\,, \quad \eta = \eta_0[x - u(y)t, y]\,. \tag{14}$$

We can alternatively view (12) in terms of Eulerian variables and in this case we are confronted with the partial differential equations

$$\frac{\partial \xi}{\partial t} + u(y)\frac{\partial \xi}{\partial x} = \eta u'(y)\,, \qquad \frac{\partial \eta}{\partial t} + u(y)\frac{\partial \eta}{\partial x} = 0\,. \tag{15}$$

It is easy to check by substitution that (14) is the solution of (15).

### 1.3 Motion is equivalent to mapping

We obtained (2) using the geometric argument in figure 1. Now we admire some different scenery by taking an algebraic path to (2). Our itinerary emphasizes that the solutions of (6) define a mapping of the space $\boldsymbol{x}_0$ of initial coordinates onto the space $\boldsymbol{x}$, and hence the title of this section.

Using indicial notation (summation implied over repeated indices), it follows from the chain rule that

$$\mathrm{d}\boldsymbol{x}_i = \frac{\partial \boldsymbol{x}_i}{\partial \boldsymbol{x}_{0j}} \mathrm{d}\boldsymbol{x}_{0j}. \tag{16}$$

Taking the time derivative of (16), and keeping in mind that $\boldsymbol{x}_{0j}$ is independent of $t$, gives

$$\frac{D}{Dt}(\mathrm{d}\boldsymbol{x}_i) = \frac{\partial \boldsymbol{u}_i}{\partial \boldsymbol{x}_{0j}} \mathrm{d}\boldsymbol{x}_{0j} = \frac{\partial \boldsymbol{u}_i}{\partial \boldsymbol{x}_{0j}} \frac{\partial \boldsymbol{x}_{0j}}{\partial \boldsymbol{x}_k} \mathrm{d}\boldsymbol{x}_k = \frac{\partial \boldsymbol{u}_i}{\partial \boldsymbol{x}_j} \mathrm{d}\boldsymbol{x}_j. \tag{17}$$

(We have used the golden rule.) Making the identification $\mathrm{d}\boldsymbol{x} \to \boldsymbol{\xi}$ we obtain (2).

The motion of a fluid defines a family of mappings from the space of initial coordinates, $\boldsymbol{x}_0$, onto the space of coordinates $\boldsymbol{x}$. At $t = 0$ this is just the identity map but as $t$ increases the map from $\boldsymbol{x}_0$ to $\boldsymbol{x}$ can become very complicated. Equation (16) defines the Jacobian matrix,

$$\mathcal{J}_{ij} \equiv \frac{\partial \boldsymbol{x}_i}{\partial \boldsymbol{x}_{0j}}, \tag{18}$$

of the map.

With these algebraic formalities we have given an alternative derivation of (2) and, as a bonus, we have also found a representation of the solution:

$$\boldsymbol{\xi} = \mathcal{J} \boldsymbol{\xi}_0. \tag{19}$$

The solution above is known as *Cauchy's* solution.

In (19) there is no assumption that the flow is incompressible. If the flow is incompressible (i.e., if $\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$) then mapping from $\boldsymbol{x}_0$ to $\boldsymbol{x}$ conserves volume. In this case, $\det \mathcal{J} = 1$.

## 2 Two-dimensional incompressible flow

In the case of a two-dimensional incompressible flow there is a streamfunction $\psi = \psi(\boldsymbol{x}, t)$ such that $\boldsymbol{u} = (u, v) = (-\psi_y, \psi_x)$. In terms of $\psi$, (2) can be written as:

$$\frac{D\boldsymbol{\xi}}{Dt} = \boldsymbol{W}\boldsymbol{\xi}, \qquad \text{where} \qquad \boldsymbol{W} \equiv \begin{pmatrix} -\psi_{xy} & -\psi_{yy} \\ \psi_{xx} & \psi_{xy} \end{pmatrix}. \tag{20}$$

The trace of $\boldsymbol{W}$ is zero and the determinant is $\det(\boldsymbol{W}) = \psi_{xx}\psi_{yy} - \psi_{xy}^2$. The solution of (20) can be written as

$$\boldsymbol{\xi} = \exp\left( \int_0^t \boldsymbol{W}(t')\, \mathrm{d}t' \right) \boldsymbol{\xi}_0 . \tag{21}$$

Thus, using (19), we obtain a fundamental connection between $\mathcal{J}(t)$ and $\boldsymbol{W}(t)$:

$$\mathcal{J}(t) = \exp\left( \int_0^t \boldsymbol{W}(t')\, \mathrm{d}t' \right). \tag{22}$$

Because $\mathrm{tr}\, \boldsymbol{W} = 0$ it follows[1] that $\det \mathcal{J} = 1$. This is, of course, just another way of saying that if the flow is incompressible then the map from $\boldsymbol{x}_0$ to $\boldsymbol{x}$ is area preserving.

[1] For a square matrix $M$

$$\det \mathrm{e}^M = \mathrm{e}^{\mathrm{tr}\, M}.$$

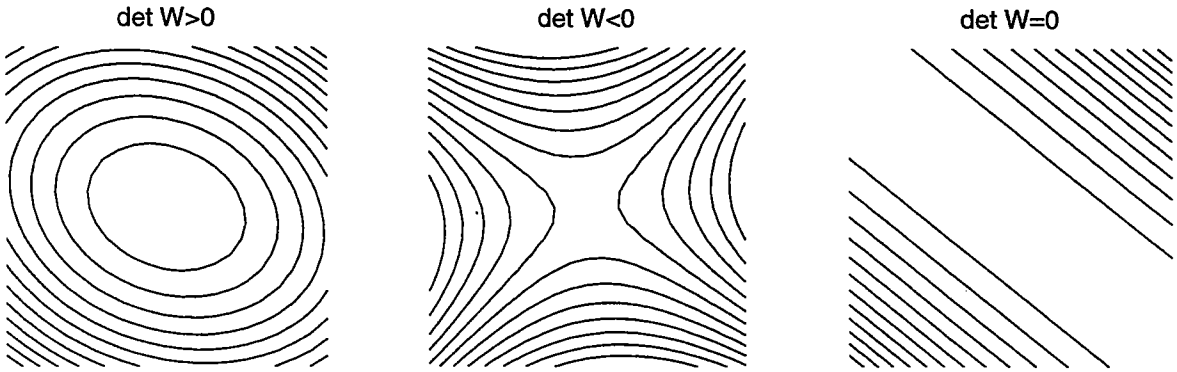


Figure 2: The sign of $\det(\boldsymbol{W}) = \psi_{xx}\psi_{yy} - \psi_{xy}^2$ determines the streamline pattern.

## 2.1 The steady case

Because (20) is linear the solution is straightforward if the velocity field in the Lagrangian frame is *steady*. Thus

$$\boldsymbol{\xi}(t) = \mathrm{e}^{\gamma t}\hat{\boldsymbol{\xi}}, \qquad \Rightarrow \qquad \gamma = \pm\sqrt{-\det \boldsymbol{W}}\,, \tag{23}$$

where

$$\det \boldsymbol{W} = \psi_{xx}\psi_{yy} - \psi_{xy}^2\,. \tag{24}$$

There are three cases, which correspond to the three panels in figure 2:

***Elliptic:*** If $\det \boldsymbol{W} > 0$, then $\gamma$ is imaginary and the local streamfunction has elliptic streamlines; $\boldsymbol{\xi}$ changes periodically in time and there is no exponential stretching.

***Hyperbolic:*** If $\det \boldsymbol{W} < 0$ then $\gamma$ is real and the streamfunction is locally hyperbolic. Then, as in lecture 1, material line elements will be stretched exponentially in one direction and compressed in the other.

***Transitional:*** If $\det \boldsymbol{W} = 0$ then $|\boldsymbol{\xi}|$ grows linearly with time.

Following Okubo (1970) and Weiss (1991), the sign of $\det \boldsymbol{W}$ has been used to diagnose two-dimensional turbulence simulations (e.g., McWilliams 1984). Assuming that $\det \boldsymbol{W}$ is changing slowly in the Lagrangian frame, one argues that the result in (23) applies "quasistatically". For instance, using simulations of two-dimensional turbulence, McWilliams shows that in the core of a strong vortex $\psi_{xx}\psi_{yy} - \psi_{xy}^2 > 0$. The interpretation is that there is no exponential stretching of line elements in vortex cores, which indicates that these regions are isolated patches of laminar flow. This so-called Okubo–Weiss criterion is only a rough guide to the stretching properties of complicated flows; for a critique and more refined results see Hua and Klein (1999).

One pleasant aspect of the steady two-dimensional case is that it is possible to explicitly calculate the matrix exponential $\boldsymbol{\mathcal{J}}(t) = \exp(t\boldsymbol{W})$. (This is not the case in three dimensions.) Begin by noting that

$$\boldsymbol{W}^2 + (\det \boldsymbol{W})\boldsymbol{\mathcal{I}} = 0\,, \tag{25}$$

where $\boldsymbol{\mathcal{I}}$ is the $2 \times 2$ identity matrix. The result above is easily checked by direct evaluation, but (25) is also a consequence of $\mathrm{tr}\,\boldsymbol{W} = 0$ and the Cayley-Hamilton theorem. When (25) is substituted into

the definition of the matrix exponential:

$$\mathcal{J} = \exp(t\boldsymbol{W}) = \mathcal{I} + t\boldsymbol{W} + \frac{t^2}{2}\boldsymbol{W}^2 + \frac{t^3}{6}\boldsymbol{W}^3 + \cdots \tag{26}$$

the sum collapses to

$$\mathcal{J} = \cos\left(\sqrt{\det \boldsymbol{W}}\, t\right)\mathcal{I} + \frac{\sin\left(\sqrt{\det \boldsymbol{W}}\, t\right)}{\sqrt{\det \boldsymbol{W}}}\boldsymbol{W} . \tag{27}$$

We now use the result above to formulate a renovation model.

## 2.2 The $\sigma$-$\zeta$ model

We construct the "$\sigma$-$\zeta$" model using the matrix equation in (20). The idea is to define an ensemble of stretching flows in which the $2 \times 2$ matrix $\boldsymbol{W}$ is piecewise constant in the intervals $\mathcal{I}_n = \{t : (n-1)\tau < t < n\tau\}$; $\tau$ is the "decorrelation time". We use the following representation of $\boldsymbol{W}$ in the interval $\mathcal{I}_n$:

$$\boldsymbol{W}_n = \frac{\zeta_n}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \frac{\sigma_n}{2}\begin{pmatrix} -\cos 2\theta_n & \sin 2\theta_n \\ \sin 2\theta_n & \cos 2\theta_n \end{pmatrix} . \tag{28}$$

$\zeta_n$ is the vorticity and $\sigma_n$ the strain. Isotropy is ensured by picking the random angle $0 < \theta_n < 2\pi$ from a uniform density. (We use $2\theta_n$ because the principal strain axes are at angle $\theta_n$ to the coordinate axes, and tbey specify a *orientation* but not an direction. That is, the strain axes are like vectors without an arrow.)

Because $\boldsymbol{W}_n$ is constant in $\mathcal{I}_n$ the calculation of stretching rates can be reduced to a product of random matrices. The terms in the product are $\exp(\tau\boldsymbol{W}_n)$ and, using (27), one can obtain this matrix exponential analytically. There is an extensive and difficult literature devoted to calculating the statistical properties of products of random matrices (e.g., Crisanti, Paladin & Vulpiani, 1993). It is fortunate that we can avoid these complications by using the isotropy of the $\sigma$-$\zeta$ model to reduce averages of matrix products to averages of scalar products.

Two important properties of $\boldsymbol{W}_n$ are easily related to the vorticity and the strain:

$$\det \boldsymbol{W}_n = \frac{1}{4}\left(\zeta_n^2 - \sigma_n^2\right) , \qquad \mathrm{tr}\left(\boldsymbol{W}_n^{\mathrm{T}}\boldsymbol{W}_n\right) = \frac{1}{2}\left(\zeta_n^2 + \sigma_n^2\right) . \tag{29}$$

In the examples that follow we will use $\sigma$-$\zeta$ ensembles which model spatially homogeneous flows, for which $\langle\sigma^2\rangle = \langle\zeta^2\rangle$ (by the way: this is not obvious). In this case $\langle \det \boldsymbol{W}_n\rangle = 0$ and "on average" the Okubo-Weiss criterion is zero.

We employ (27) to obtain an explicit expression for the matrix $\mathcal{J}_n = \exp(\tau\boldsymbol{W}_n)$. It turns out that we do not need the full details: all that is required is

$$\frac{1}{2}\mathrm{tr}\left(\mathcal{J}_n^{\mathrm{T}}\mathcal{J}_n\right) = 1 + \Xi(\sigma_n, \tau_n, \tau) , \tag{30}$$

where

$$\Xi(\sigma, \zeta, \tau) \equiv \frac{\sigma^2}{\zeta^2 - \sigma^2}\left[1 - \cos\left(\sqrt{\zeta^2 - \sigma^2}\,\tau\right)\right] . \tag{31}$$

The "trace formula" above should be known to experts on two-dimensional stretching problems, but I have not found (30) in the literature.

### 2.2.1 Stretching of squared length

Consider the first interval $\mathcal{I}_1$, and suppose that at $t = 0$, $\boldsymbol{\xi} = \ell_0(\cos\chi, \sin\chi)$. At $t = \tau$ we have

$$\ell_1^2 = \boldsymbol{\xi}_0^{\mathrm{T}} \boldsymbol{\mathcal{J}}_1{}^{\mathrm{T}} \boldsymbol{\mathcal{J}}_1 \boldsymbol{\xi}_0 \,. \tag{32}$$

Now we use isotropy to average (32) over the random direction $\chi$ of the element $\boldsymbol{\xi}_0$. A trivial calculation gives

$$\langle (\ell_1/\ell_0)^2 \rangle_\chi = \frac{1}{2} \mathrm{tr} \left( \boldsymbol{\mathcal{J}}_1{}^{\mathrm{T}} \boldsymbol{\mathcal{J}}_1 \right) . \tag{33}$$

The RHS of (33) is given explicitly in (30). We must still average over the random variables $\sigma$ and $\zeta$. This gives

$$\langle (\ell_1/\ell_0)^2 \rangle = 1 + \int\!\!\int \mathcal{P}(\sigma,\zeta) \Xi(\sigma,\zeta,\tau) \, \mathrm{d}\sigma \mathrm{d}\zeta \,, \tag{34}$$

where $\mathcal{P}(\sigma,\zeta)$ is the joint PDF of $\sigma$ and $\zeta$ [2].

We are now well on our way to computing the rate at which $\ell^2$ grows with the number of renovation cycles, $n$. The average stretching of $\ell^2$ in each $\mathcal{I}_n$ is independent of the previous $\mathcal{I}$'s. Thus, to compute the growth of $\ell^2$ over $n$ renovation cycles, we can simply raise the average $\ell^2$-stretching factor in a single $\mathcal{I}$ to the $n$'th power:

$$\langle (\ell_n/\ell_0)^2 \rangle = \left\{ 1 + \int\!\!\int \mathcal{P}(\sigma,\zeta) \Xi(\sigma,\zeta,\tau) \, \mathrm{d}\sigma \mathrm{d}\zeta \right\}^n . \tag{35}$$

The stretching rate $\gamma_2$ is defined by

$$\gamma_2 \equiv \lim_{t\to\infty} \frac{1}{t} \ln \left[ \langle (\ell_n/\ell_0)^2 \rangle^{1/2} \right] . \tag{36}$$

The notation $\gamma_2$ anticipates section 4 in which we will define a stretching rate $\gamma_p$ which measures the growth of $\langle (\ell_n/\ell_0)^p \rangle$.

Using $n = t/\tau$, we have from (35)

$$\gamma_2 = \frac{1}{2\tau} \ln \left\{ 1 + \int\!\!\int \mathcal{P}(\sigma,\zeta) \Xi(\sigma,\zeta,\tau) \, \mathrm{d}\sigma \mathrm{d}\zeta \right\} . \tag{37}$$

To further simplify the integral above we must specify the probability density function $\mathcal{P}(\sigma,\zeta)$ (examples follow).

### 2.2.2 Randomly oriented Couette flows

As a first example, suppose that in each $\mathcal{I}_n$ the random variables $\zeta_n$ and $\sigma_n$ are independently and identically distributed, each equal to $\pm\beta$ with equal probability. In this case

$$\mathcal{P}(\sigma,\zeta) = \frac{1}{4} \left[ \delta(\sigma+\beta) + \delta(\sigma-\beta) \right] \left[ \delta(\zeta+\beta) + \delta(\zeta-\beta) \right] . \tag{38}$$

This ensemble is a set of randomly oriented Couette flow, such as the third case in figure 2. According to the Okubo-Weiss criterion there should be no stretching because det $\boldsymbol{W}$ is identically zero. However, this is wrong.

[2] If $\sigma$ and $\zeta$ are independent and identically distributed random varaibles then $\mathcal{P}(\sigma,\zeta) = \hat{\mathcal{P}}(\sigma)\hat{\mathcal{P}}(\zeta)$.

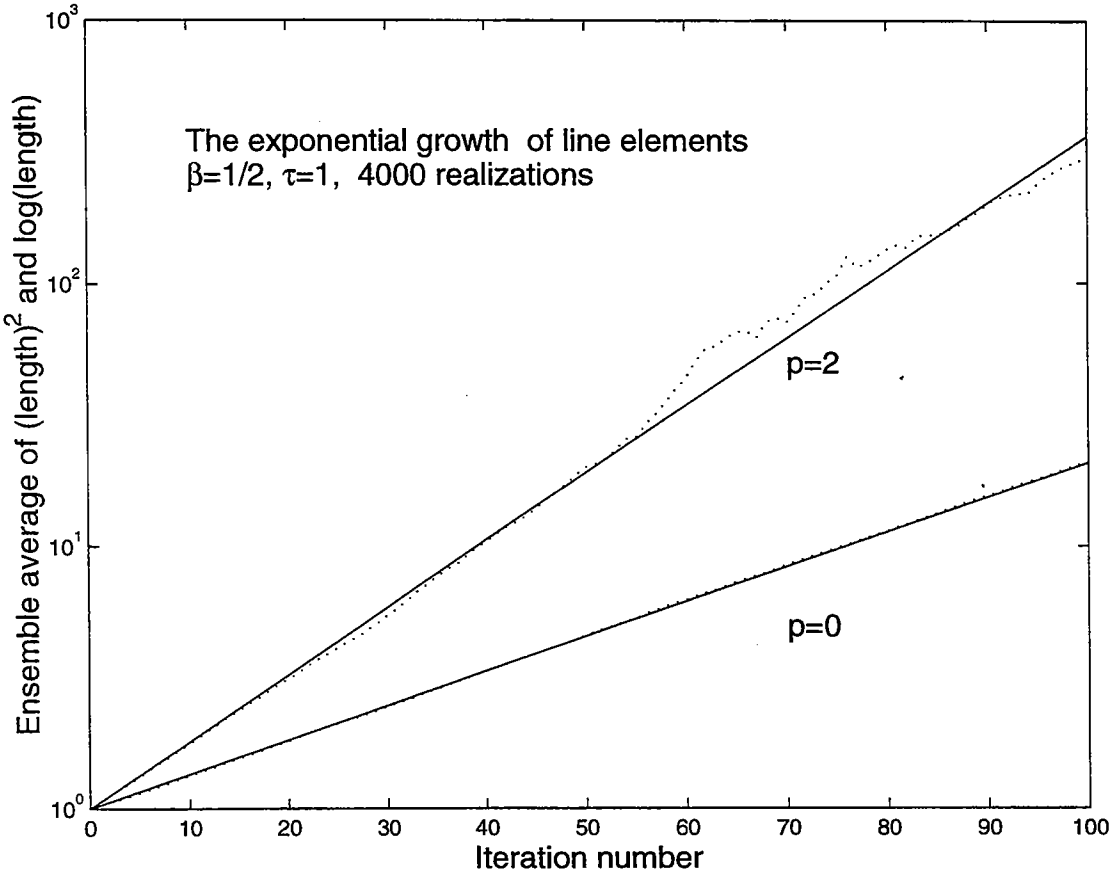


Figure 3: A comparison of the exponent $\gamma_2$ in (40) with a simulation (the dotted curves) of the random Couette flow. The simulation is conducted by creating random matrices according to the recipe in (28) and (38), and then computing the matrix product. The iteration number is the number of matrices in the product. To get reasonable agreement between the simulation and the analytic result in (40) one must ensemble average over a large number of realizations (4000 in the figure above). The discrepancies evident at large iteration number can be reduced by using more realizations. The figure also shows a comparison of the exponent $\gamma_0$ in (73) with simulation.

The recipe in (38) leads to trivial calculations because $\zeta_n^2 = \sigma_n^2$ and $\Xi = \beta^2\tau^2/2$. Thus, even without averaging over $\sigma$ and $\zeta$,

$$\frac{1}{2}\mathrm{tr}\left(\mathcal{J}_n^{\mathrm{T}}\mathcal{J}_n\right) = 1 + \frac{1}{2}\beta^2\tau^2\,, \tag{39}$$

and it follows that

$$\gamma_2 = \frac{1}{2\tau}\ln\left(1 + \frac{\beta^2\tau^2}{2}\right)\,. \tag{40}$$

(See Figure 3.) The nonzero exponential stretching, which occurs even though $\det \boldsymbol{W} = 0$, is due to the realignment of a material element with respect to the direction of extension of the velocity field which occurs at $t = n\tau$. In the limit of a very slowly changing velocity field, $\tau \to \infty$, the stretching rate vanishes because there are fewer realignment events. This is the revenge of the Okubo-Weiss criterion.

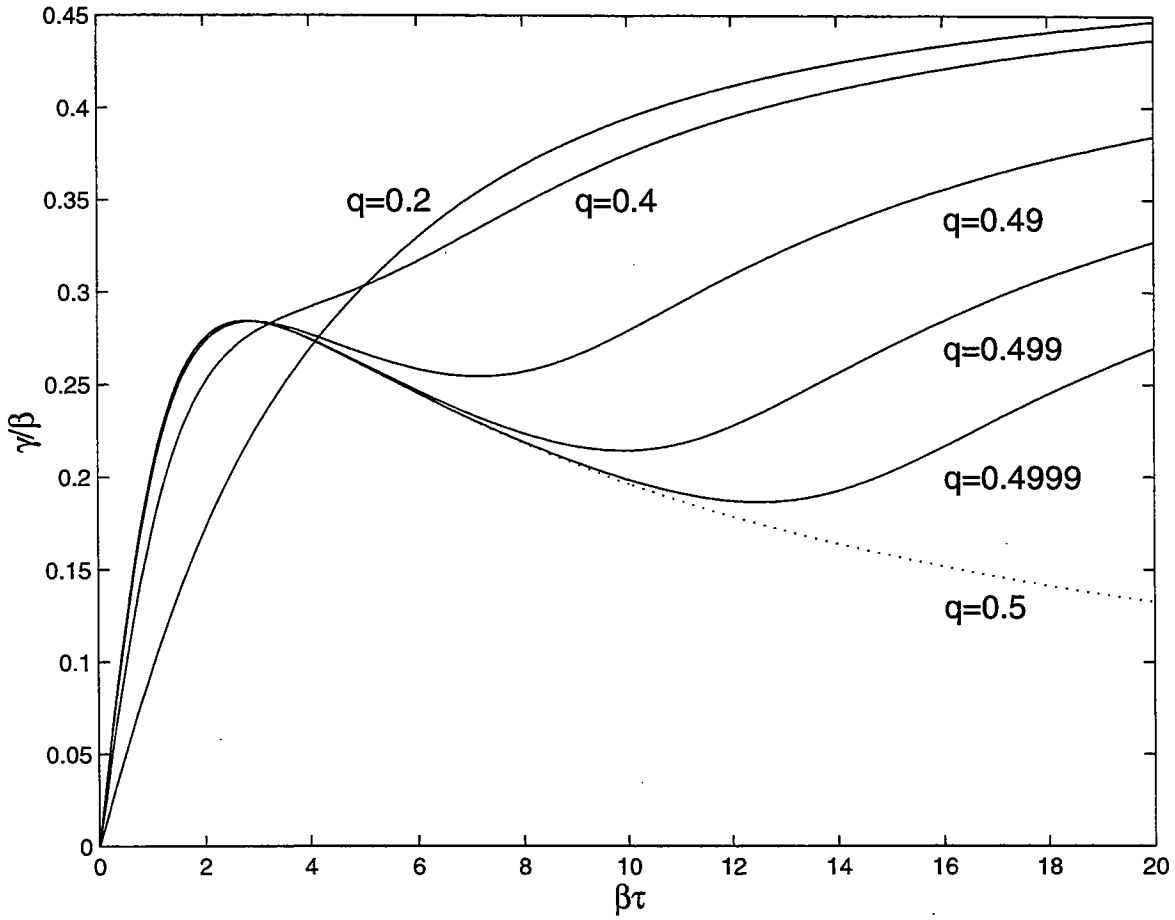


Figure 4: The nondimensional stretching exponents $\gamma_2/\beta$ in (41) as a function of $\beta\tau$ for various values of $q$. If $q = 1/2$, then $\det \boldsymbol{W}$ is zero identically and $\gamma \to 0$ as $\tau \to \infty$. When slightly less than $1/2$, and $\tau$ is sufficiently large, the occasional hyperbolic points can make a large contribution to the stretching exponent $\gamma_2$.

### 2.2.3 An example with $\det \boldsymbol{W} \neq 0$

A more interesting stretching ensemble is defined by taking $\sigma_n$ and $\zeta_n$ to be identical and independently distributed random variables equal to $\beta$ with probability $q$, $-\beta$ with probability $q$, or zero with probability $1 - 2q$. With this prescription there is a hyperbolic point in $\mathcal{I}_n$, as in the middle panel of figure 2, with probability $2q(1 - 2q)$.

One can calculate $\gamma_2$ in (37) by enumeration and averaging over the nine possible pairs $(\sigma_n, \zeta_n)$. Calculation gives

$$\gamma_2 = \frac{1}{2\tau} \ln \left\{ 1 + 2q^2\beta^2\tau^2 + 2q(1-2q)\left(\cosh \beta\tau - 1\right) \right\} . \tag{41}$$

Figure 4 shows the nondimensional $\gamma_2/\beta$ as a function of $\beta\tau$ for various values of $q$. From figure 4 we conclude that while instantaneous hyperbolic points are not essential for exponential stretching, they do help, especially if the correlation time $\tau$ is long.

### 2.2.4 The Batchelor and Kraichnan limits

The calculation of stretching exponents in this section does not follow the historical path. The pioneering papers by Batchelor (1959) and Kraichnan (1974) considered limiting cases — slowly decorrelating in the case of Batchelor and rapidly decorrelating in the case of Kraichnan — in which stretching rates can be calculated approximately. A major advantage of these approximations is that they work equally well in two and three dimensional space. On the other hand, by considering exactly soluble two-dimensional models we can extract the Batchelor and Kraichnan limits as special cases.

Batchelor (1959) considered stretching by slowly decorrelating velocity fields. This is the limit in which $\zeta\tau$ and $\sigma\tau$ are large. Batchelor's main conclusion is that in this quasisteady limit the net stretching is dominated by hyperbolic straining events. Indeed, this conclusion is illustrated by the exact result for $\gamma_2$ which is plotted in figure 4.

Kraichnan (1974) considered the opposite limit in which $\zeta\tau$ and $\sigma\tau$ are small. In this rapidly decorrelating limit we can simplify the exact expression in (37) by noting that $\Xi \approx (\sigma\tau)^2/2 \ll 1$. Thus, simplifying (37), we find that the stretching rate is

$$\gamma_2 \approx \frac{1}{4}\langle\sigma^2\rangle\tau\,, \tag{42}$$

independent of the vorticity.

## 2.3 The renovating wave model

In this section we calculate the average growth of $\ell^2$ using the renovating wave (RW) model. It is interesting to see how this calculation can be done without using matrix identities such as (27).

Begin by recalling the definition of the RW model. The RW streamfunction is

$$\mathcal{I}_n = (n-1)\tau_* < t < n\tau_*: \qquad \psi_n \equiv \cos\left[\cos\theta_n\, x + \sin\theta_n\, y + \varphi_n\right]. \tag{43}$$

In (43), $\theta_n$ and $\varphi_n$ are random phases and $\tau_*$ is the decorrelation time. The random phases are reinitialized at $t = n\tau_*$ so there is the complete and sudden loss of memory at these instants. (In this section we use the dimensionless version of the RW model; the parameter $\tau_* \equiv \tau kU$ is the ratio of the correlation time $\tau$ to the maximum shear of the sinusoidal wave $kU$.)

The renovating wave model is equivalent to the random map

$$(x_{n+1}, y_{n+1}) = (x_n, y_n) + (s_n, -c_n)\sin[c_n x_n + s_n y_n + \varphi_n]\tau_*\,, \tag{44}$$

where $(c_n, s_n) \equiv (\cos\theta_n, \sin\theta_n)$. The Jacobian matrix can easily be obtained by differentiation of (44):

$$\mathcal{J}^{(n)} = \exp(\tau_* \boldsymbol{W}^{(n)}) = \begin{bmatrix} 1 + c_n s_n \tau_* \psi_n & s_n^2 \tau_* \psi_n \\ -c_n^2 \tau_* \psi_n & 1 - c_n s_n \tau_* \psi_n \end{bmatrix}. \tag{45}$$

Notice that $\det \mathcal{J}^{(n)} = 1$: the map is area preserving.

Using $\mathcal{J}^{(n)}$ we can track the stretching of an infinitesimal material element as

$$\boldsymbol{\xi}_{n+1} = \mathcal{J}^{(n)}\boldsymbol{\xi}_n, \qquad \Rightarrow \qquad \ell^2_{n+1} = \boldsymbol{\xi}^{\mathrm{T}}_{n+1}\boldsymbol{\xi}_{n+1} = \boldsymbol{\xi}^{\mathrm{T}}_n \mathcal{K}^{(n)} \boldsymbol{\xi}_n, \tag{46}$$

where $\mathcal{K}^{(n)} = \mathcal{J}^{(n)\mathrm{T}}\mathcal{J}^{(n)}$. Explicitly:

$$\mathcal{K}^{(n)} = \begin{bmatrix} (1 + c_n s_n \psi_n \tau_*)^2 + c_n^4 \psi_n^2 \tau_*^2 & (s_n^2 - c_n^2)\psi_n \tau_* + c_n s_n \psi_n^2 \tau_*^2 \\ (s_n^2 - c_n^2)\psi_n \tau_* + c_n s_n \psi_n^2 \tau_*^2 & (1 - c_n s_n \psi_n \tau_*)^2 + s_n^4 \psi_n^2 \tau_*^2 \end{bmatrix}. \tag{47}$$

To compute the stretching rate we consider an element which has length $\ell_0$ at $t = 0$. Because the problem is isotropic, it is harmless to choose the coordinate system so that this element lies along the $x$-axis: $\boldsymbol{\xi}_0 = \ell_0(1, 0)$. After the first iteration of the map:

$$\ell_1^2 = \mathcal{K}_{11}^{(1)} \ell_0^2 = \left[(1 + c_1 s_1 \psi_1 \tau_*)^2 + c_1^4 \psi_1^2 \tau_*^2\right] \ell_0^2. \tag{48}$$

Averaging (48) over the phases $\theta_1$ and $\varphi_1$ gives

$$\langle (\ell_1/\ell_0)^2 \rangle = \left(1 + \frac{\tau_*^2}{4}\right). \tag{49}$$

If you are suspicious of the argument above, then you might prefer to align the initial material element at an arbitrary angle, say $\boldsymbol{\xi}_0 = \ell_0(\cos\chi, \sin\chi)$. Repeating the calculation, we now find that

$$\ell_1^2 = \left(\mathcal{K}_{11}^{(1)} \cos^2\chi + \mathcal{K}_{22}^{(1)} \sin^2\chi\right) \ell_0^2, \tag{50}$$

Averaging (50) over $\theta_1$ and $\varphi_1$, we recover (49).

Because each $\mathcal{J}^{(n)}$ is independent of the earlier $\mathcal{J}$'s the average growth of $\ell^2$ is

$$\langle (\ell_n/\ell_0)^2 \rangle = \left(1 + \frac{\tau_*^2}{4}\right)^n. \tag{51}$$

Using $t = n\tau_*$, (51) can be written as

$$\langle (\ell_n/\ell_0)^2 \rangle^{1/2} = \mathrm{e}^{\gamma_2 t}, \qquad \gamma_2 \equiv \frac{1}{2\tau_*} \ln\left(1 + \frac{\tau_*^2}{4}\right). \tag{52}$$

Aside from notational differences, the expression above for $\gamma_2$ is identical to (40).

# 3 Amplification of concentration gradients

In this section we discuss the amplification of $\boldsymbol{\nabla} c$ which occurs when a passive scalar is advected by a random flow in two dimensions.

Back in (4) we noted that the quantity $\boldsymbol{\xi} \cdot \boldsymbol{\nabla} c$ satisfies the conservation equation

$$\frac{D}{Dt}(\boldsymbol{\xi} \cdot \boldsymbol{\nabla} c) = 0. \tag{53}$$

Equation (53) enables us to use our earlier results concerning the stretching of material elements to analyze gradient amplification. In fact, using (53), we can obtain $\boldsymbol{\nabla} c$ from $\boldsymbol{\xi}$. The first step is to construct a basis by considering the following initial value problem:

$$\frac{D\boldsymbol{\xi}_k}{Dt} = (\boldsymbol{\xi}_k \cdot \boldsymbol{\nabla}) \boldsymbol{u}, \qquad \text{with initial conditions} \quad \boldsymbol{\xi}_1(\boldsymbol{x}, 0) = \hat{\boldsymbol{x}}, \quad \boldsymbol{\xi}_2(\boldsymbol{x}, 0) = \hat{\boldsymbol{y}}, \tag{54}$$

where the unit vectors of the coordinate system are $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}$. As the fluid moves, the parallelogram spanned by $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ will deform. But because $\boldsymbol{u}$ is incompressible, the area of the parallelogram is constant and so

$$\boldsymbol{\xi}_1 \times \boldsymbol{\xi}_2 = \hat{\boldsymbol{z}}, \qquad (\text{for all } t). \tag{55}$$

If we can solve (54) for $\boldsymbol{\xi}_1$, then we can use (53) and (55) to calculate $\boldsymbol{\xi}_2$ and $\boldsymbol{\nabla} c$.

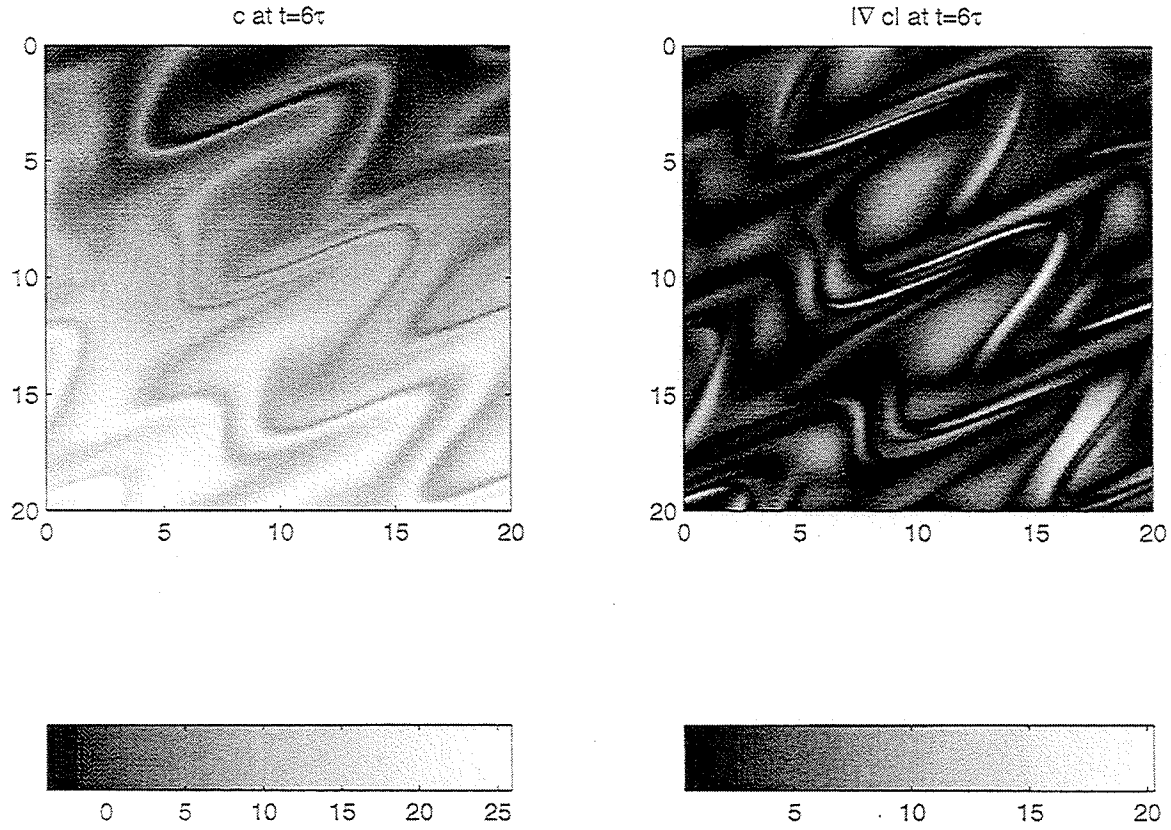


Figure 5: Numerical solution of the renovating wave model with $\tau = 2$. The initial condition is $c(x, y, 0) = y$. Already, at $t = 6\tau$, $|\nabla c|$ is greatly amplified in some regions.

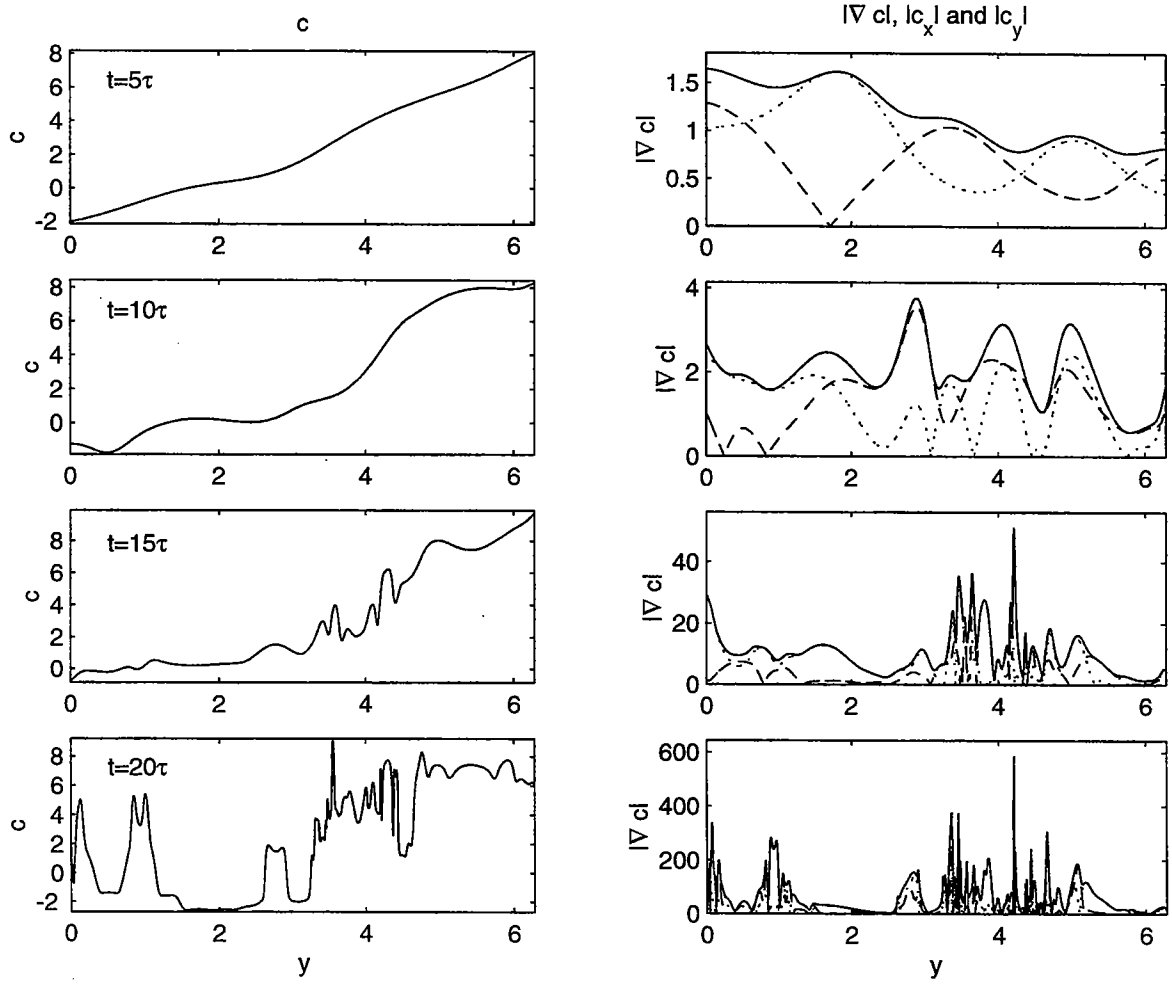


Figure 6: A numerical solution of the renovating wave model with $\tau = 1$. The initial condition is $c(x, y, 0) = y$. The plots show the values of $c$ and $|\nabla c|$ along the slice $x = 0$. After 20 iterations, $|\nabla c|$ has developed strong spatial intermittency.

As an example of this procedure, suppose that the initial condition is $c(\boldsymbol{x},0)=y$. Then it follows from (53) that:

$$\boldsymbol{\xi}_1\cdot\boldsymbol{\nabla}c=0 \qquad \text{and} \qquad \boldsymbol{\xi}_2\cdot\boldsymbol{\nabla}c=1 \qquad (\text{for all } t). \tag{56}$$

Using (55) and (56) we see that

$$\boldsymbol{\nabla}c=\hat{\boldsymbol{z}}\times\boldsymbol{\xi}_1. \tag{57}$$

Thus, in this example, once we calculate $\boldsymbol{\xi}_1$ we obtain $\boldsymbol{\nabla}c$ as a bonus.

Figure 5 displays the numerical solution for $c$ and $|\boldsymbol{\nabla}c|$ after 6 iterations of the renovating wave model. The initial condition is $c(\boldsymbol{x},0)=y$, so that $\boldsymbol{\nabla}c(\boldsymbol{x},0)=\hat{\boldsymbol{y}}$; the decorrelation time is $\tau=2$. The field in figure 5 is obtained using a $256\times256$ grid. To find $c$ at the grid point $\boldsymbol{x}$ at time $t=n\tau$, one iterates the renovating wave model backwards in time till the initial location $(a,b)$ is determined, and then $c(\boldsymbol{x},t)=b$. In parallel with this backwards iteration, $\boldsymbol{\xi}(\boldsymbol{x},n\tau)$ is computed by matrix multiplication of the $\mathcal{J}^{(n)}$ defined in (45), and then $\boldsymbol{\nabla}c$ is given by (57).

An important feature of stirring is the development of *intermittency* in the concentration gradient, $|\boldsymbol{\nabla}c|$. In figure 6 the development of intermittency is illustrated, again using the renovating wave model. After 20 iterations there are "hotspots" in which large values of $|\boldsymbol{\nabla}c|$ are concentrated. Without diffusion, the gradient of $c$ condenses onto a fractal set as the number of iterations increases (Városi, Antonsen & Ott 1991).

# 4 Multiplicative random variables

In our solution of the $\sigma$-$\zeta$ model in section 2 we used isotropy to reduce a product of random matrices to a product of random scalars e.g., see equations (33), (34) and the following discussion. The main point of this section is that the statistical properties of isotropic line-element stretching are bedevilled by the large fluctuations which are characteristic of products of random variables. Indeed, figure 6 shows that there are large fluctuations in $\ell^2=|\boldsymbol{\xi}|^2=|\boldsymbol{\nabla}c|^2$. If one is attempting to measure the variance dissipation, $\kappa\langle\boldsymbol{\nabla}c\cdot\boldsymbol{\nabla}c\rangle$, then the intermittent structure of $\boldsymbol{\nabla}c$ in figure 5 might pose a sampling problem. Imagine steering a ship through the field in figure 5 and making occasional point measurements of $\boldsymbol{\nabla}c$. If the density of measurements is too low then one might easily miss the gradient hot-spots and so grossly underestimate $\kappa\langle\boldsymbol{\nabla}c\cdot\boldsymbol{\nabla}c\rangle$.

## 4.1 Most probable values versus mean values

We begin by stepping back from the stirring problem, and making some general remarks about multiplicative random processes. Suppose that a random quantity, $X$, is formed by taking the product of $N$ independent and identically distributed random variables

$$X=x_1x_2\cdots x_N. \tag{58}$$

What can we say about the statistical properties of $X$?

The most nonintuitive aspect of $X$ in (58) is the crucial distinction which must be made between the *mean* value of $X$ and the *most probable* value of $X$. As an illustration, it is useful to consider an extreme case in which each $x_k$ in (58) is either $x_k=0$ or $x_k=2$ with equal probability. Then the sample space consists of $2^N$ sequences of zeros and two's. For all but one those sequences, $X=0$; in the remaining single case $X=2^N$. Thus, the most probable (that is, most frequently occuring) value of $X$ is

$$X_{\rm mp}=0\,. \tag{59}$$

On the other hand, the mean of $X$ is

$$\langle X\rangle \equiv \frac{\text{sum all the } X\text{'s from different realizations}}{\text{number of realizations}} = 1 \tag{60}$$

Notice that one can also calculate $\langle X\rangle$ by arguing that $\langle x_k\rangle = 1$ and, since the $x_k$'s are independent, $\langle X\rangle = \langle x_k\rangle^N = 1$

The example above is representative of multiplicative processes in that extreme events, although exponentially rare if $N \gg 1$, are exponentially different from typical or most probable events. Thus, for the *product* of $N$ random variables the ratio $\langle X\rangle/X_{\rm mp}$ diverges exponentially as $N \to \infty$. On the other hand, for the *sum* of $N$ random variables the most probable outcome is a good approximation of the mean outcome. Perhaps this is why people have an intuitive appreciation of sums, but find products confusing.

Now let us consider a more realistic example in which each $x_k$ is either $\alpha$ or $1/\alpha$ with probability $1/2$. In this case the $p$'th moment of $X$ is

$$\langle x_k^p\rangle = \frac{1}{2}\left(\alpha^p + \alpha^{-p}\right), \qquad \Rightarrow \qquad \langle X^p\rangle = \left(\frac{\alpha^p + \alpha^{-p}}{2}\right)^N. \tag{61}$$

Before continuing, the student will profit from showing that the most probable value of $X$ is $X_{\rm mp} = 1$ (for $N$ even). For example, if $\alpha = 2$ then $\langle X\rangle = (5/4)^N$, while $X_{\rm mp} = 1$. Again, the most probable value differs exponentially from the mean value as $N \to \infty$.

## 4.2 The log-normal distribution

Because $X_{\rm mp}$ is so different from the $\langle X\rangle$ the problem of determining $\langle X\rangle$ via Monte Carlo simulation is difficult: one may have to exhaust nearly all of the $2^N$ cases in order to obtain a reliable estimate of $\langle X\rangle$. This exhaustion is necessary for the first example, in which $x_k = 0$ or 2. In the example of equation (61), provided that $\alpha \approx 1$, we can get a pretty good estimate of $\langle X\rangle$ with less than exhaustive enumeration of all sequences of the $x_n$'s.

Begin by noting that

$$\ln X = \ln x_1 + \ln x_2 + \cdots + \ln x_N, \tag{62}$$

and so if $\ln x_k$ has finite variance then it follows from the Central Limit Theorem (CLT) that $\Lambda \equiv \ln X$ becomes normally distributed as $N \to \infty$.

The pitfall is in concluding that all the important statistical properties of $\Lambda$, and therefore of $X = \exp(\Lambda)$, can be calculated using the asymptotic log-normal distribution of $X$. This not the case because the PDF of $\Lambda$, $\mathcal{P}(\Lambda)$, is approximated by a Gaussian only in a central scaling region in which $|\Lambda| < cN^{1/2}$, where $c$ is some constant which depends on the PDF of $x_k$. On the other hand, a reliable calculation of $\langle X^p\rangle = \langle \exp(p\Lambda)\rangle$ may require knowledge of the tail-structure of $\mathcal{P}(\Lambda)$.

To illustrate these difficulties, we use the example in which $\ln x_k = \pm \ln\alpha$ and $\langle \ln^2 x_k\rangle = \ln^2\alpha$. Invoking the Central Limit Theorem, the asymptotic PDF of $\Lambda$ is therefore

$$\mathcal{P}_{\rm CLT}(\Lambda) = \frac{1}{\sqrt{2\pi N \ln^2\alpha}} \exp\left(-\Lambda^2/2N\ln^2\alpha\right). \tag{63}$$

In the central scaling region, $\mathcal{P}(\Lambda) \approx \mathcal{P}_{\rm CLT}(\Lambda)$.

To determine $X_{\rm mp}$ we can consider $\Lambda = \ln X$, which is an additive process for which the mean and most probable coincide, so that

$$\langle \ln X\rangle = \ln X_{\rm mp}, \qquad \Rightarrow \qquad X_{\rm mp} = \mathrm{e}^{\langle \ln X\rangle}. \tag{64}$$

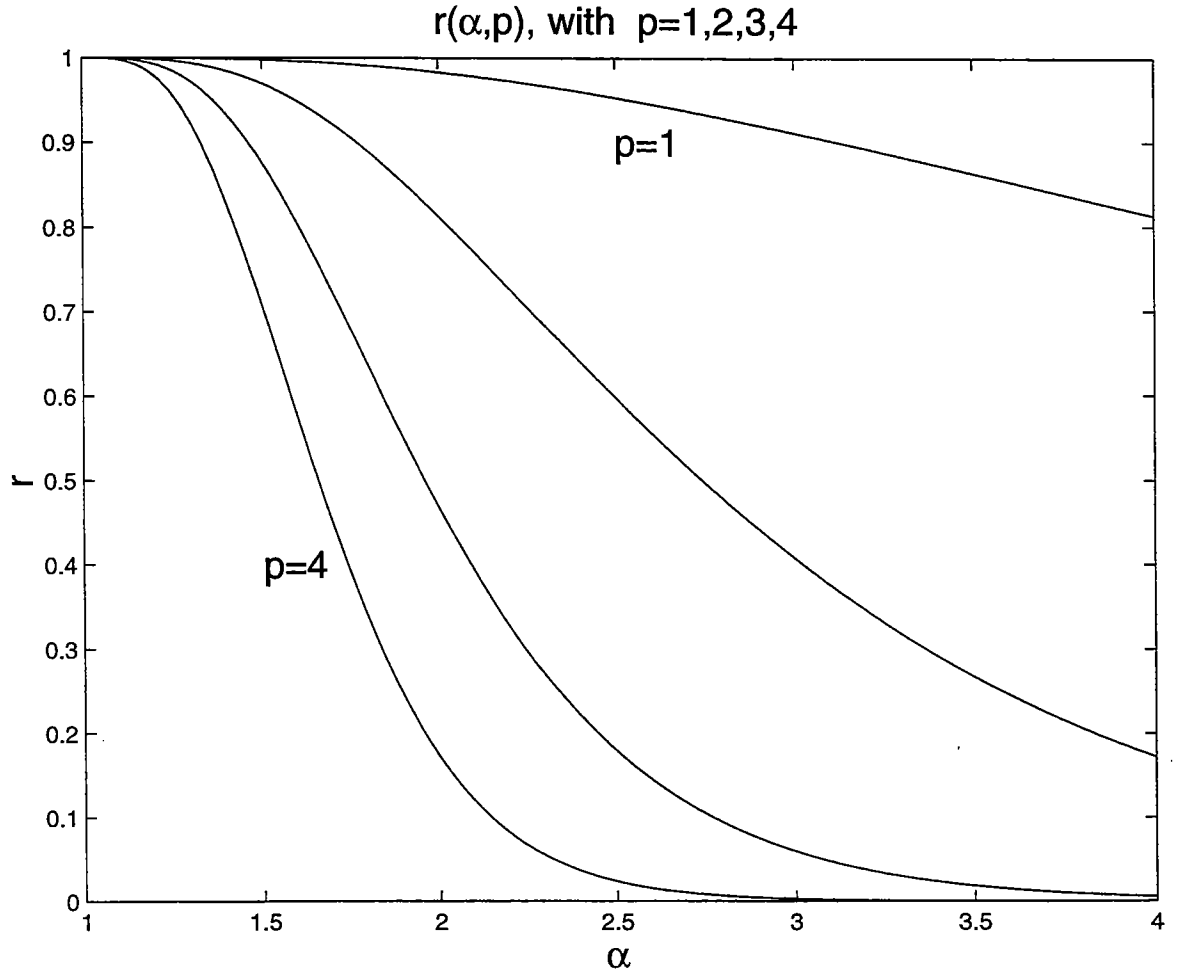


Figure 7: The function $r(\alpha, p)$ defined in (67). In order to accurately estimate $\langle X^p \rangle$ using the CLT one must have $r \approx 1$.

In our example with $\ln x_k = \pm \ln \alpha$, $\langle \ln X \rangle = 0$ and therefore $X_{\rm mp} = 1$. (This is one way of solving the problem posed in the previous section; another is to obtain the exact $\mathcal{P}(\Lambda)$ using the binomial density.)

With hope in our hearts, we now attempt to recover the exact result in (61) by substituting (63) into

$$\langle X^p \rangle \equiv \int_{-\infty}^{\infty} \mathrm{e}^{p\Lambda} \mathcal{P}(\Lambda)\, \mathrm{d}\Lambda. \tag{65}$$

After the integration, one finds that

$$\langle X^p \rangle_{\rm CLT} = \exp\left(N p^2 \ln^2 \alpha / 2\right). \tag{66}$$

To assess the error we form the ratio of the exact result to the approximation:

$$\langle X^p \rangle / \langle X^p \rangle_{\rm CLT} = r^N, \quad \text{where} \quad r \equiv \frac{1}{2} \exp\left(-p^2 \ln^2 \alpha / 2\right) \left(\alpha^p + \alpha^{-p}\right). \tag{67}$$

When $r(\alpha,p)$ is close to 1, the error is tolerable in the sense that $\ln\langle X^p\rangle_{\rm CLT}$ is close to $\ln\langle X^p\rangle$.

For example, with $\alpha = 2$, the exact result is $\langle X\rangle = (5/4)^N$ while $\langle X\rangle_{\rm CLT} = (1.27)^N$. However the second moment $p = 2$, is seriously in error. As a general rule, $\langle X^p\rangle_{\rm CLT}$ is a reliable estimate of $\langle X^p\rangle$ provided that $p^2\langle \ln^2 x_k\rangle < c$, where $c$ is the constant which determines the width of central scaling region, $|\Lambda| < cN^{1/2}$, in which $\mathcal{P}(\Lambda) \approx \mathcal{P}_{\rm CLT}(\Lambda)$. We conclude that the complete analysis of a random multiplicative quantity cannot be reduced to the Central Limit Theorem merely by taking a logarithm.

## 4.3 Stretching exponents

Equation (64) is a very important result for multiplicative random variables: *to obtain the most probable value of $X$, one can exponentiate $\langle \ln X\rangle$*. This explains why there is so much attention paid to $\langle \ln[\ell(t)/\ell_0]\rangle$ in the literature on random line element stretching: knowing the average of the logarithm enables one to estimate the stretching of a *typical* line element. Of course, the typical line element may not make a large contribution to the dissipation $\kappa\langle \nabla c'\cdot\nabla c'\rangle$. Thus our earlier focus on $\ell^2$ was not wasted, but it was not complete either.

A good characterization of random stretching is provided by the complete set of *stretching exponents*. Following Drummond & Münch, we define the stretching exponents, $\gamma_p$, as

$$\gamma_p \equiv \lim_{t\to\infty} \frac{1}{p\langle \ell^p\rangle}\frac{{\rm d}\langle \ell^p\rangle}{{\rm d}t}, \qquad p>0, \tag{68}$$

and

$$\gamma_0 = \lim_{p\to 0}\gamma_p = \lim_{t\to\infty}\frac{\rm d}{{\rm d}t}\langle \ln \ell\rangle . \tag{69}$$

Knowing all these $\gamma$'s, the asymptotic growth of line elements is characterized by

$$\langle \ell^p\rangle^{1/p} \sim \ell_0 {\rm e}^{\gamma_p t}. \tag{70}$$

Back in section 2 we calculated only $\gamma_2$ (e.g., see (33) and the subsequent discussion). To conclude this section I will discuss the calculation of the other stretching exponents, particularly $\gamma_0$.

## 4.4 The stretching exponent $\gamma_0$ of the $\sigma$-$\zeta$ model

As an example of the difference between $\gamma_0$ and $\gamma_2$ we return to the $\sigma$-$\zeta$ model. In section 2 we obtained a general expression for $\gamma_2$ in (37). Now consider the problem of determining $\gamma_0$. Taking the log of (32), writing $\boldsymbol{\xi}_0 = \ell_0(\cos\chi, \sin\chi)$, and then integrating[3] over $\chi$, we have after some travail,

$$\langle \ln(\ell_1/\ell_0)\rangle_\chi = \frac{1}{2}\ln\left(1+\frac{\Xi}{2}\right), \tag{71}$$

where $\Xi(\sigma,\zeta,\tau)$ is given in (31). Averaging over $\sigma$ and $\zeta$, and using $\gamma_0 = \tau^{-1}\langle \ln(\ell_1/\ell_0)\rangle$, gives

$$\gamma_0 = \frac{1}{2\tau}\int\!\!\int \mathcal{P}(\sigma,\zeta)\ln\left[1+\frac{1}{2}\Xi(\sigma,\zeta,\tau)\right]{\rm d}\sigma{\rm d}\zeta . \tag{72}$$

The expression above should be compared with that for $\gamma_2$ in (37).

[3]The integral

$$\int_0^\pi \ln(a\pm b\cos x)\,{\rm d}x = \pi\ln\left[\left(a+\sqrt{a^2-b^2}\right)/2\right],$$

is useful.

With the ensemble of random Couette flows in section 2.2.2 we can evaluate the integrals in (37) and (72). Thus, we find that

$$\gamma_0 = \frac{1}{2\tau}\ln\left(1+\frac{\beta^2\tau^2}{4}\right), \quad \gamma_2 = \frac{1}{2\tau}\ln\left(1+\frac{\beta^2\tau^2}{2}\right). \tag{73}$$

Notice that $\gamma_2 > \gamma_0$. This is a illustration of the general result that $\gamma_p$ is an increasing function of $p$ (Childress & Gilbert 1995). Figure 3 compares the expressions for $\gamma_0$ and $\gamma_2$ in (73) with simulation.

## 4.5 Stretching in one-dimension

One-dimensional compressible velocity fields provide striking examples of the nontrivial dependence of $\gamma_p$ on $p$. We conclude this lecture with a model of random one-dimensional stretching for which the $\gamma_p$'s can be obtained analytically.

### 4.5.1 A sinusoidal velocity

With the one-dimensional velocity $u = \sin x$, the equation governing line element stretching, (2), is

$$\xi_t + \sin x\,\xi_x = \xi\cos x\,, \quad \xi(x,0) = 1\,. \tag{74}$$

The initial condition above is that the line elements attached to different fluid particles all have the same initial length. Because the fluid is compressible, the fluid density $\rho(x,t)$ satisfies

$$\rho_t + (\sin x\,\rho)_x = 0 \qquad \rho(x,0) = 1. \tag{75}$$

It is easy to show by substitution that the solutions of (74) and (75) are related $\rho(x,t) = 1/\xi(x,t)$. The physical interpretation of this result should be obvious...

To solve (74), we follow the route outlined in section 1.3 by determining the mapping from the initial space, $x_0$, to the space $x(x_0,t)$. This means we solve

$$\frac{Dx}{Dt} = \sin x, \qquad x(0,x_0) = x_0\,. \tag{76}$$

Using separation of variables we find that

$$\tan(x/2) = \mathrm{e}^t\tan(x_0/2)\,, \tag{77}$$

which enables us to determine $x$ given $x_0$, or vice versa. Figure 8 shows how the mapping from $x_0$ to $x$ evolves as $t$ increases.

In this one-dimensional example, the Jacobian of the mapping is simply

$$\frac{\mathrm{d}x}{\mathrm{d}x_0} = \frac{1}{\cosh t - \cos x_0\sinh t} = \cosh t + \cos x\sinh t\,. \tag{78}$$

It is easy to check by substitution that $\xi = \mathrm{d}x/\mathrm{d}x_0$ is the solution of (74).

### 4.5.2 A one-dimensional renovating model

Using the previous one-dimensional illustration of the Cauchy solution, we can formulate a renovating model that illustrates some of the subtleties involved in random stretching problems. Consider an ensemble of random renovating one-dimensional velocity fields in which

$$u = \sin(x+\varphi_n) \qquad \text{if} \qquad (n-1)\tau < t < n\tau\,. \tag{79}$$

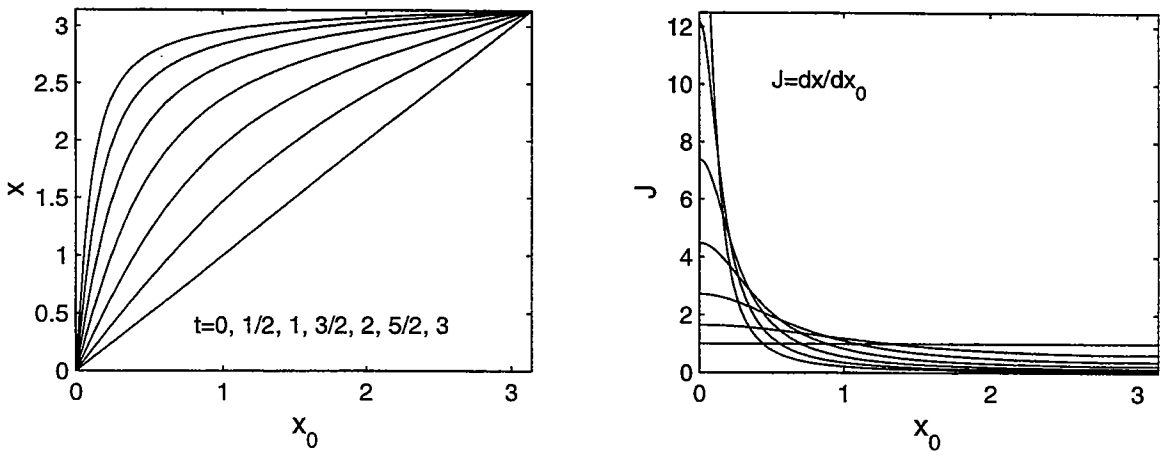


Figure 8: The left panel shows the mapping from $x_0$ to $x$ at the indicated times. The interval $0 < x_0 < \pi$ is compressed into the neighbourhood of $x = \pi$. The right panel shows $J(x_0, t)$ at the same times. Notice that an element that starts at say, $x_0 = 1/2$, is first stretched ($J > 1$) but then ultimately compressed ($J < 1$) as the particle approaches $x = \pi$.

The random phase, $0 < \varphi_n < 2\pi$, is reset at $t = n\tau$. Notice that there is no preferred location on the $x$-axis; that is, the statistical properties of the process are spatially homogeneous.

Now, suppose we follow the stretching of a line element attached to a particle that moves in a particular realization of this velocity field. We denote location of this particle at $t = n\tau$ by $a_n$, and the length of the attached line element at this time by $\ell_n$. Then, using the solution from the previous section, the stretching of the line element is given by the random product

$$\ell_n = J(a_{n-1})J(a_{n-2})\cdots J(a_0)\ell_0\,, \tag{80}$$

where the Jacobian is

$$J(a) \equiv \frac{1}{\cosh\tau - \cos a \sinh\tau}\,. \tag{81}$$

Because the phase is reset at $t = n\tau$, each $J(a_n)$ in (80) is independent of the others. Moreover, because of spatial homogeneity, each $a_n$ is uniformly distributed with $0 < a_n < 2\pi$.

Equation (80) expresses the length of a material line element at $t = n\tau$ as a product of $n$ random numbers. Following our earlier discussion of multiplicative random variables, we first calculate $\gamma_0$ by taking the logarithm of (80):

$$\ln(\ell_n/\ell_0) = \sum_{k=0}^{n-1} \ln J(a_k)\,, \tag{82}$$

Thus, the mean of $\ln(\ell_n/\ell_0)$ is

$$\langle \ln(\ell_n/\ell_0)\rangle = n\langle \ln J\rangle\,, \tag{83}$$

where

$$\langle \ln J\rangle = \oint \ln\left[J(a)\right] \frac{\mathrm{d}a}{2\pi} = -\ln\left[\cosh(\tau/2)\right]\,. \tag{84}$$

Because $\langle(\ln J)^2\rangle$ is finite, the central limit theorem applies and we conclude that as $n \to \infty$, $\ln(\ell_n/\ell_0)$ is approximately normally distributed with the mean value $n\langle \ln J\rangle$.

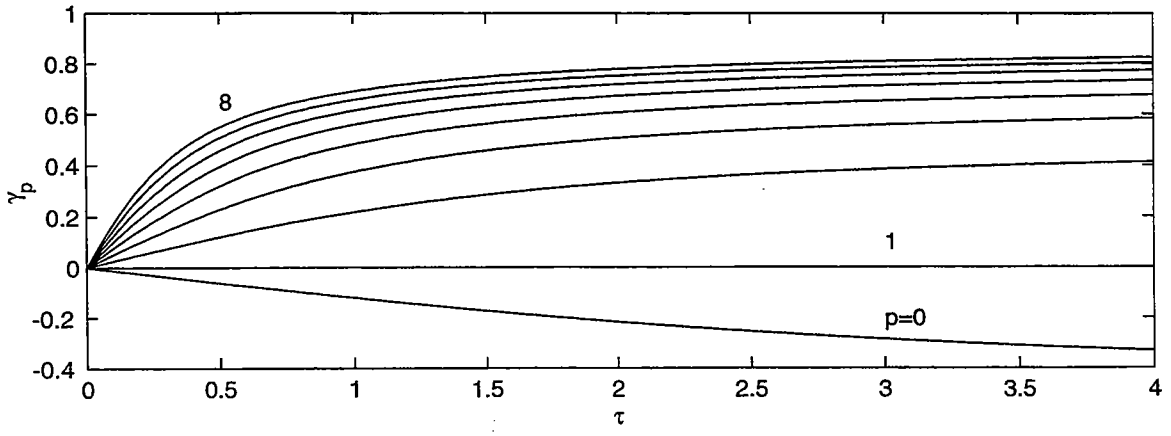


Figure 9: The stretching exponents $\gamma_p(\tau)$, with $p = 0, 1, \cdots, 8$ calculated using (89) .

Moreover, we can conclude from the central limit theorem that the *most probable* value of $\ell_n/\ell_0$ is

$$(\ell_n/\ell_0)_{\rm mp} \approx \mathrm{e}^{\langle \ln(\ell_n/\ell_0)\rangle} = \mathrm{e}^{\gamma_0 t}\,, \tag{85}$$

where, since $n = t/\tau$,

$$\gamma_0 = -\ln[\cosh(\tau/2)]/\tau < 0\,. \tag{86}$$

The result in (85) is remarkable because it implies that most of the line elements in this compressible flow exponentially contract (rather than stretch) as $t \to \infty$!

Exponential contraction of most material lines is incomplete disagreement with the spirit of Batchelor's result in (1), where $\gamma > 0$. The result above, that $\gamma_0 < 0$, is a special consequence of the *compressible* velocity field used in (79). (For a discussion of compressible velocities in a space of arbitrary dimension, see Chertkov et al. (1998).) In any event, this example shows that one cannot take exponential stretching for granted, no matter how intuitive it seems on the basis of experiments, such as those of Welander (1955).

How is contraction in the length of most material elements compatible with conservation of the total length of the $x$-axis? The answer is that even though most elements become exponentially small as $t \to \infty$, a few elements become exponentially large. Thus most of the length accumulates in exponentially rare, but exponentially long, line elements. This is an elementary example of an *inverse cascade* i.e., the spontaneous appearence of large-scale structures (big line elements).

To demonstrate length conservation, we can compute the mean (as opposed the most probable) length of an element. The mean length is

$$\langle \ell_n \rangle = \langle J \rangle^n \ell_0\,, \tag{87}$$

where $J(a)$ is defined in (81) and

$$\langle J \rangle = \oint J(a)\frac{\mathrm{d}a}{2\pi} = 1\,. \tag{88}$$

Thus, the mean length of an element is constant, even though most elements exponentially contract.

As an exercise, I suggest showing that for integer values of $p$ the stretching exponents of this one-dimensional model are given by

$$\gamma_p = \ln\left[P_{p-1}(\cosh\tau)\right]/p\tau\,, \tag{89}$$

where $P_m$ is the $m$'th Legendre polynomial. Thus, in this particular example, there is a nice analytic characterization of the rate at which different moments stretch (see figure 9).

# References


[1] A.Okubo. Horizontal dispersion of floatable particles in the vicinity of velocity singularity such as convergences. *Deep-Sea Res.*, 17:445–454, 1970.

[2] G. K. Batchelor. Small–scale variation of convected quantities like temperature in turbulent fluid. part 1. general discussion and the case of small conductivity. *J. Fluid Mech.*, 5:113–133, 1959.

[3] G.K. Batchelor. The effect of turbulence on material lines and surfaces. *Proc. Roy. Soc. London A*, 213:349–366, 1952.

[4] B.L.Hua and P. Klein. An exact criterion for the stirring properties of nearly two-dimensional turbulence. *Physica D*, 113:98–110, 1999.

[5] M. Chertkov, I. Kolokolov, and M. Vergassola. Inverse versus direct cascades in turbulent advection. *Phys. Rev. Lett.*, 80(3):512–515, 1998.

[6] S. Childress and A. D. Gilbert. *Stretch, Twist, Fold: The Fast Dynamo.* Springer, Berlin, 1995.

[7] W.J. Cocke. Turbulent hydrodynamic line stretching: consequences of isotropy. *Phys. Fluids*, 12:2488–2492, 1969.

[8] A. Crisanti, G. Paladin, and A. Vulpiani. *Products of Random Matrices.* Springer-Verlag, Berlin, 1993.

[9] I.T. Drummond and W. Münch. Turbulent stretching of line and surface elements. *J. Fluid Mech.*, 215:45–59, January 1990.

[10] R.H. Kraichnan. Convection of a passive scalar by a quasi-uniform random stretching field. *J. Fluid Mech.*, 64:737–762, 1974.

[11] J.C. McWilliams. The emergence of isolated coherent vortices in turbulent flows. *J. Fluid Mech.*, 198:199–230, 1984.

[12] S.A. Orszag. Comments on 'turbulent hydrodynamic line stretching: consequences of isotropy. *Phys. Fluids*, 13:2203–2204, 1970.

[13] F. Varosi, T.M. Antonsen, and E. Ott. The spectrum of fractal dimensions of passively convected scalar gradients. *Phys. Fluids A*, 3:1017–1028, 1991.

[14] J. Weiss. The dynamics of enstrophy transfer in two-dimensional hydrodynamics. *Physica D*, 48:273–294, 1991.

[15] P. Welander. Studies on the general development of motion in a two-dimensional, ideal fluid. *Tellus*, 7:141–156, 1955.

# Lecture 4: Anomalous diffusion

In this lecture we discuss stochastic models of correlated random walks. By "correlated" we mean that that if a particle is headed in one direction then there is nonzero probability that it continues in that same direction for some time and this probability fades to zero as the time interval increases. This is, of course, the situation envisaged by Taylor (1921).

The distinction between normal and anomalous diffusion made in lecture 1 can be understood by examining the rate at which velocity correlation decrease to zero. Normal diffusion occurs if the velocity correlation decrease rapidly while anomalous diffusion results from processes in which particles move coherently for long times with infrequent changes of direction. Roughly speaking, this distinction is quantified by the tail behaviour of the velocity autocorrelation function. For example, if the correlation function decays exponentially then there is normal diffusion, whereas if the correlation function decays algebraically then there is the possibility of anomalous diffusion.

The definition of anomalous diffusion is based only on the behaviour of the second moment, $\langle x^2 \rangle$. But we usually want to know more about the distribution of a tracer than simply the second moment. In the case of normal diffusion, detailed information concerning the tracer distribution is obtained by solving the diffusion equation

$$c_t = Dc_{xx} \,. \tag{1}$$

Can we obtain continuum models, analogous to (1), which provide the same detailed information for anomalously diffusing tracer? The main goal of this lecture is to develop partial differential equation models which can be used for this purpose.

## 1 Superdiffusion and subdiffusion

### 1.1 Taylor's formula and long tails

Yet again we recall Taylor's formula which relates the growth of position variance to an integral of the Lagrangian velocity autocorrelation function, $\mathrm{corr}(t)$,

$$\frac{\mathrm{d}\langle x^2 \rangle}{\mathrm{d}t} = 2 \int_0^t \mathrm{corr}(t') \, \mathrm{d}t' \,. \tag{2}$$

In order to obtain $\langle x^2 \rangle$ we must integrate (2). Standard manipulations turn the resulting double integral of corr$(t)$ into a single integral

$$\langle x^2 \rangle = 2 \int_0^t (t - t') \text{corr}(t') \, dt' . \tag{3}$$

The result (3), which is not in Taylor's original paper, will prove to be very useful.

We usually have in mind situations in which corr$(t)$ decreases to zero as $t \to \infty$ so that the integrals in (2) and (3) converge to nonzero values. An example is the renovating wave model, with its "triangular" correlation function, from lecture 2. Later in this lecture I will introduce the telegraph model which has an exponentially decaying correlation function, corr$(t) = U^2 \exp(-2\alpha t)$. These are both examples in which correlations decrease very rapidly so that normal diffusion occurs. But now consider the possibility that corr$(t)$ decreases so slowly that the integrals in (3) diverge.

Suppose, for instance, that as $t \to \infty$, corr$(t) \sim t^{-\eta}$ with $0 < \eta < 1$. Even though the diffusivity no longer exists, it still follows from (3) that

$$\langle x^2 \rangle \sim t^{2-\eta} . \tag{4}$$

In this case there is *superdiffusion*: the variance of the particle displacement grows faster than linearly with time because $2 - \eta > 1$.

Taylor's formula also contains the possibility of *subdiffusion*. This case is subtle because, like the example of the sea-surface mentioned in lecture 1, it requires that the integral defining $D$ is zero. But suppose additionally that the remaining integral in (3) diverges. This can happen if corr$(t) \sim ct^{-\eta}$ with $1 < \eta < 2$. The condition that $1 < \eta$ ensures that $\int_0^\infty \text{corr}(t') \, dt'$ converges (to zero). The second inequality, $\eta < 2$, ensures that $\int_0^t t' \text{corr}(t') \, dt'$ diverges. Using (3), we again find the scaling law in (4). However this time, because $2 - \eta < 1$, there is *subdiffusion*.

At first glance two possibilities above appear as unlikely exceptions to the more natural cases in which both integrals in (3) converge. However there are examples in fluid mechanics in which either subdiffusion or superdiffusion is observed experimentally or computationally. Thus (4) cannot be dismissed as an unlikely pathology.

## 1.2 The Texas experiments

An experiment illustrating anomalous diffusion has been conducted in Swinney's laboratory at University of Texas; see Solomon, Weeks & Swinney (1994) and Weeks, Urbach & Swinney (1996). These investigators study the dispersion of particles in an almost two-dimensional flow in annular tank (see figure 1). The tank is rotating at about 1 or 2 Hertz and the bottom is sloped to simulate the $\beta$-effect. Because of the rapid rotation the flow is quasi two-dimensional.

The flow is forced by pumping fluid through the tank. If the pumping rate is sufficiently large then this azimuthal flow is unstable to a vortex-forming instability.

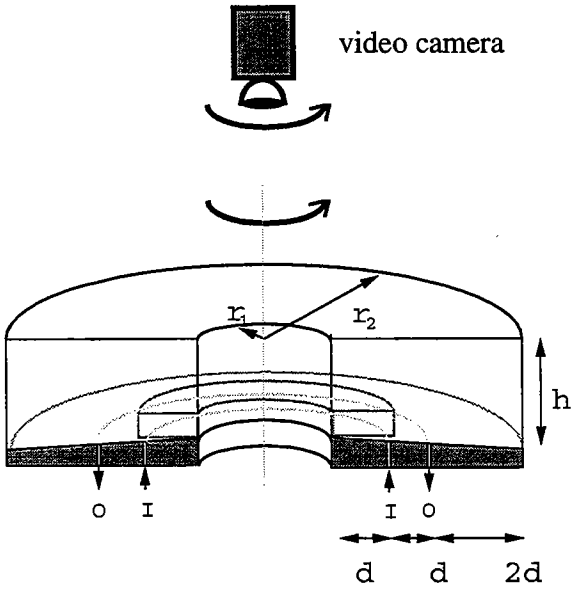


Figure 1: A sketch of the rotating annulus; the rotation rate is about 1Hz. Flow is forced by pumping water in through the ring of holes marked by I and withdrawing the same volume through the other ring marked O. As a consequence of the strong Coriolis forced acting on the radial flow between these concentric rings there is an azimuthal flow around the annulus. The experiment is viewed from above using a video camera. Figure courtesy of Eric Weeks.

A typical flow pattern in the rotating frame is shown in figure 2. Evident also in this figure is the azimuthal jet which runs all the way around the tank. The vortex pattern can be perturbed experimentally by making the strength of the pumping depend on azimuth. In this fashion, one can drive an unsteady flow and observe chaotic particle trajectories.

Automated image processing techniques are used to follow nearly neutrally buoyant tracer particles suspended in such flow. Typical particle trajectories are shown in figure 3. Particles within a vortex remain trapped for very long time (*stick*). Particles in the azimuthal jet experience prolonged *flights* around the circumference of the tank. Because the vortex pattern is not perfectly stationary particles alternate, apparently randomly, between flying in jets and sticking in vortices.

One can change the pattern of jets and vortices by altering the diameters of the circular barriers which confine the flow. Thus it is possible to create a flow with two oppositely directed jets separated by a vortex chain. In this case the dispersion process is more symmetric than in figure 4 because the flights go in both directions around the tank.

During a flight the angular displacement is proportional to the time elapsed since

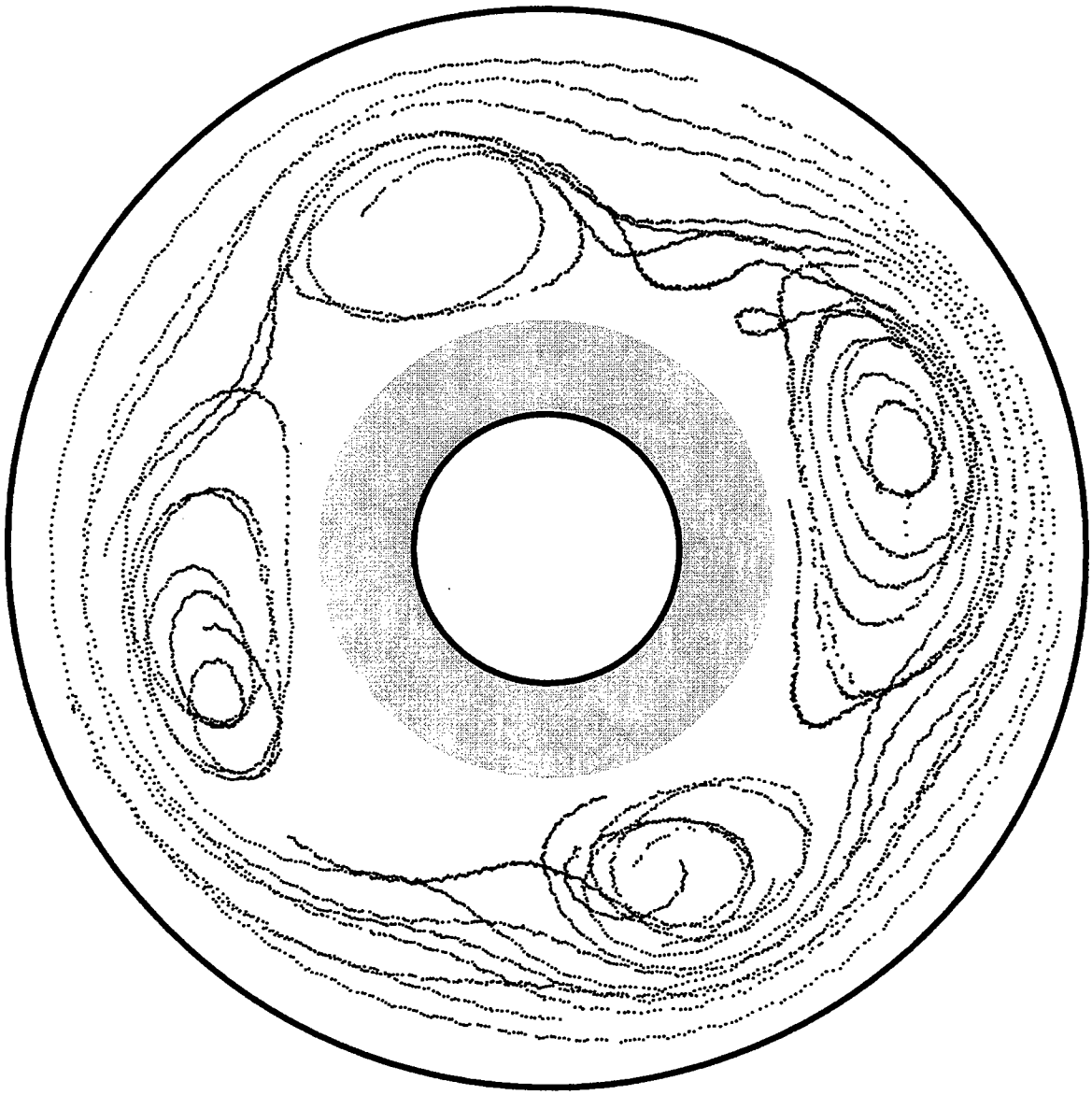

Figure 2: Streaks formed by 100s trajectories of 12 particles reveal four vortices. Weeks et al. show that the motion of these coherent vortices is chaotic. That is, a velocity spectrum, obtained by measuring velocity with a hot film probe, is broad band. Figure courtesy of Eric Weeks.

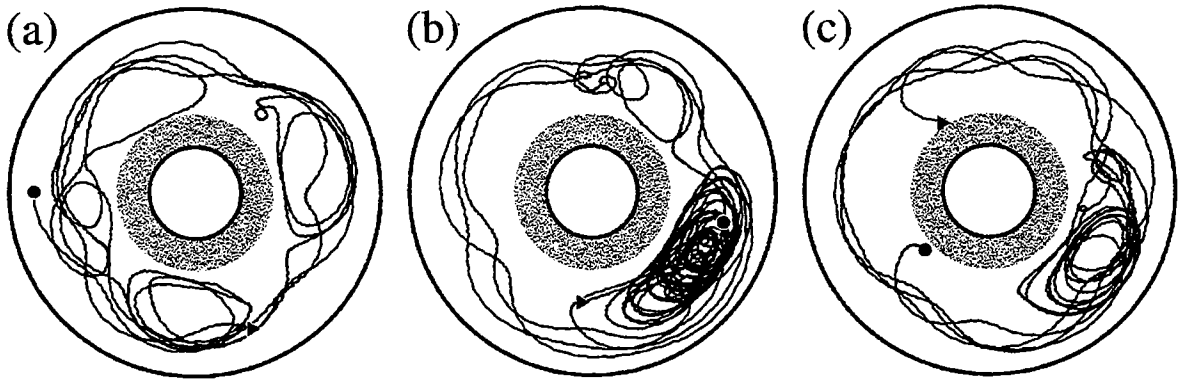


Figure 3: Trajectories of three tracer particles in the flow shown in figure 2. The beginning of each trajectory is indicated by a triangle and the end by with a circle. In (b) the particle spends most of its life trapped in a single vortex. However, this vortex wobbles erratically because the flow is chaotic. In parts (a) and (c) the particles experience several episodes of trapping within a vortex and flight around the tank in the jet. Figure courtesy of Eric Weeks.

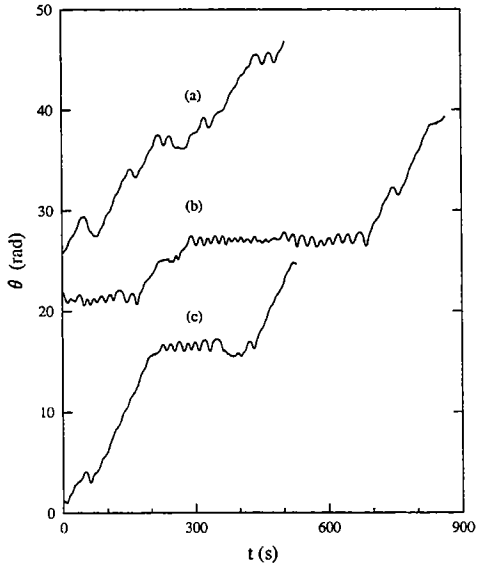


Figure 4: Angular displacement, $\theta(t)$ for the trajectories in figure 3. There is an obvious distinction between the flights and the sticking events. The small oscillations during the sticking events correspond to particle motion within a vortex. Figure courtesy of Eric Weeks.

the flight began:

$$\Delta\theta \approx Ut. \tag{5}$$

The displacement, $\Delta\theta$, is essentially zero during a sticking event (see figure 4).

The experiments show that the dispersion of an ensemble of particles is superdiffusive. That is

$$\langle(\theta - \langle\theta\rangle)^2\rangle \sim t^\gamma \tag{6}$$

where $\gamma > 1$; typical values are $\gamma \approx 1.4$ to 1.7 depending on the experimental configuration. (It is also possible to observe normal diffusion, $\gamma = 1$, by strongly forcing the flow and breaking the azimuthal symmetry of the forcing.)

To characterize the motion Solomon et al. used sticking and flying PDFs:

$$\mathcal{P}_F(a)\mathrm{d}a = \text{ Probability that a flight has a duration} \in (a, a + \mathrm{d}a)\,. \tag{7}$$

Later in this lecture we will refer to $a$ as the "lifetime" of a particle in the flying or sticking state. We figuratively speak of a particle being born into the flying state and moving coherently for a lifetime $a$ so that the total angular displacement during the flight is $\Delta\theta = Ua$.

The PDF $\mathcal{P}_F$ is normalized by $\int_0^\infty \mathcal{P}_F(a)\mathrm{d}a = 1$ and

$$\tau_F = \int_0^\infty a\mathcal{P}_F(a)\,\mathrm{d}a = \text{average duration of a flight}\,. \tag{8}$$

The PDF of sticking times, $\mathcal{P}_S(a)$, and the average sticking time, $\tau_S$, are defined analogously.

Experiments show that as $a \to \infty$, $\mathcal{P}_F$ and $\mathcal{P}_S$ have algebraically decaying tails:

$$\mathcal{P}_F(a) \sim a^{-\mu_F}, \qquad \mathcal{P}_S(a) \sim a^{-\mu_S}, \tag{9}$$

with

$$2 < (\mu_F, \mu_S) < 3. \tag{10}$$

Because of this slow algebraic decay the variance of the lifetimes, defined by

$$\langle a^2\rangle_{F,S} \equiv \int_0^\infty a^2\mathcal{P}_{F,S}(a)\,\mathrm{d}a\,, \tag{11}$$

diverges.

The divergence of $\langle a^2\rangle_F$ is significant because invoking Einstein's formula for the diffusivity

$$D = \frac{\langle(\Delta\theta)^2\rangle}{\tau}\,, \tag{12}$$

and using $\Delta\theta = Ua$, we conclude that $D \propto \langle(\Delta\theta)^2\rangle = U^2\langle a^2\rangle_F = \infty$. The divergence of $D$ is symptomatic of superdiffusion.

Notice that the denominator $\tau$ in (12) is related to the average flying and sticking times, $\tau_F$ and $\tau_S$, which are both finite. Thus, in the Texas experiments, we can say that anomalous diffusion occurs because the numerator of (12) is divergent. In other cases it is the denominator which causes trouble.

The Texas experiments show that anomalous diffusion occurs in realistic and geophysically relevant systems. Several theoretical questions suggest themselves. How do the algebraic tails of $\mathcal{P}_S$ and $\mathcal{P}_F$ arise, and can we make a microscopic models which exhibits this phenomenon? Can we relate the exponents $\gamma$, $\mu_F$ and $\mu_S$? (From section 4, the answer to the last question is $\gamma = 4 - \mu_F$.)

# 2 The telegraph model

The key issue raised by anomalous diffusion is decay of velocity correlations. Thus our goal is to formulate models for which we can explicitly calculate velocity statistics and understand the decay of correlations. Our first attempt is not very ambitious: we begin with the *telegraph model*, which is the simplest example of a continuous-time correlated random walk.

## 2.1 The Lagrangian formulation of the telegraph model

In a telegraph process the velocity of particle $n$, denoted by $u_n(t)$, can have only one of two possible values, $+U$ and $-U$. The velocity of each particle, $u_n(t)$, flips randomly back and forth between $\pm U$ with a transition probability $\alpha$ per time. This means that in a time $\mathrm{d}t$ a fraction $\alpha \mathrm{d}t$ of the ensemble switches velocity. Because the transition rate, $\alpha$, is constant we can say that a particle has no "memory" of when it first arrived in its present state. Thus this telegraph model is Markovian.

We refer to the prescription for constructing a telegraph process as *model A*. There is a variant, *model B*, discussed below.

With the prescription above, the velocity of a particle is a discontinuous function of time as shown in figure 5. The correlation function and the diffusivity are

$$\mathrm{corr}(t) = U^2 e^{-2\alpha|t|}, \qquad D = \int_0^\infty \mathrm{corr}(t)\,\mathrm{d}t = \frac{U^2}{2\alpha}, \qquad \text{(model A)}. \tag{13}$$

Notice that the $\mathrm{corr}_{tt}$ is infinite at $t = 0$; this is because the acceleration is infinite at the discontinuities in figure 5.

To obtain (13), return to the definition of the correlation function

$$\mathrm{corr}(t) = \frac{1}{N}\sum_{n=1}^{N} u_n(0)u_n(t), \tag{14}$$

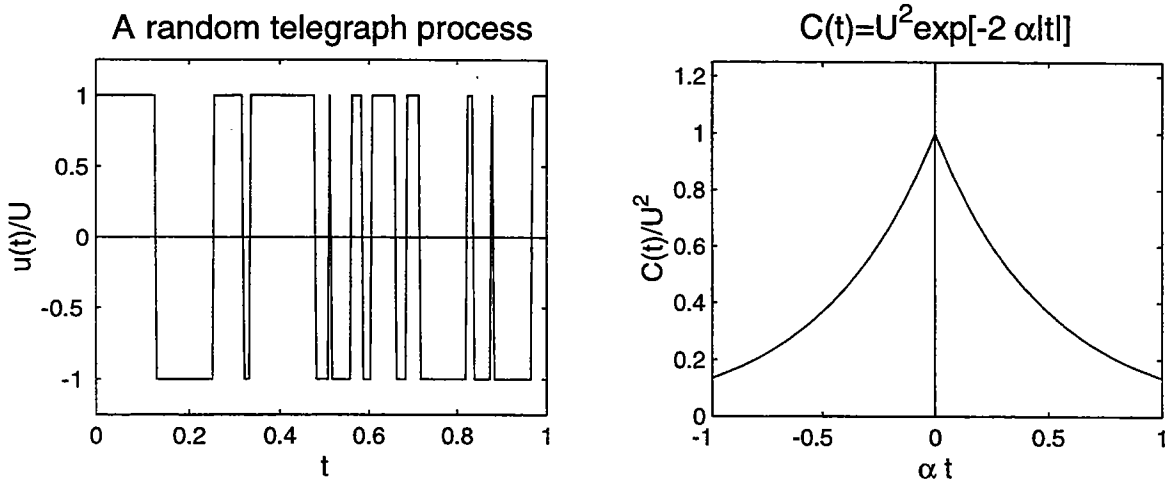


Figure 5: An example of a telegraph time series, and the telegraph correlation function.

where $N$ is the total number of particles in the ensemble. Suppose that at $t$ the sum on the right hand side has $P(t)$ positive terms, all equal to $U^2$, and $N - P(t)$ negative terms, all equal to $-U^2$. Thus

$$\mathrm{corr}(t) = \frac{U^2}{N}[2P(t) - N]\,. \tag{15}$$

In a time $\mathrm{d}t$, $P\alpha\mathrm{d}t$ of the positive terms become negative and $(N-P)\alpha\mathrm{d}t$ of the negative terms become positive. Thus, at $t + \mathrm{d}t$,

$$P(t + \mathrm{d}t) = P(t)(1 - 2\alpha\mathrm{d}t) + N\alpha\mathrm{d}t\,, \tag{16}$$

and the analog of (15) is:

$$\mathrm{corr}(t + \mathrm{d}t) = \frac{U^2}{N}[2P(t)(1 - 2\alpha\mathrm{d}t) + 2N\alpha\mathrm{d}t - N]\,. \tag{17}$$

Taking the limit $\mathrm{d}t \to 0$ in (17) gives $\mathrm{corr}_t = -2\alpha\mathrm{corr}$; the solution of this differential equation is (13).

An alternative telegraph process (model B) is constructed by imagining that at random instants each particle flips at coin. The flipping rate is $\alpha$ so that in a time $\mathrm{d}t$, there are $N\alpha\mathrm{d}t$ coin flips. After each flip, the velocity is $+U$ if there is a head and $-U$ if a tail. With this prescription, a particle will change direction on average once out of every two tosses. On the other tosses the particle continues in the same direction and the result is as if nothing happened. Thus with model B we simply replace $\alpha$ by $\alpha/2$ in our earlier calculations and consequently the correlation function and diffusivity are

$$\mathrm{corr}(t) = U^2 e^{-\alpha|t|}, \qquad D = \int_0^\infty \mathrm{corr}(t)\,\mathrm{d}t = \frac{U^2}{\alpha}, \qquad \text{(model B)}. \tag{18}$$

The difference between model A and model B is trivial. However the distinction between the two cases will plague us later.

If we are searching for a model of anomalous diffusion then the telegraph model is a disappointment: the exponentially decaying correlation function ensures that $D$ is finite and that the displacement variance ultimately grows diffusively. We continue our investigation of the telegraph model in order to better understand "ultimately" and because in section 4 the telegraph model is used as the foundation of more elaborate models which do show anomalous diffusion.

## 2.2 The Eulerian formulation of the telegraph model

Now we ignore the Lagrangian information contained in the correlation function (14) and instead we give an Eulerian formulation of the telegraph process. Let $R(x,t)$ denote the density (particles/length) of particles moving to the right with velocity $+U$ and $L(x,t)$ denote the density of left-moving particles with velocity $-U$. The coupled conservations laws are

$$R_t + UR_x = \alpha(L-R), \qquad L_t - UL_x = \alpha(R-L). \tag{19}$$

These equations should be self-evident...

We can put (19) into a revealing alternative form by defining the total concentration, $C(x,t)$, and the flux, $F(x,t)$, as

$$C \equiv R + L\,, \qquad F \equiv U(R-L)\,. \tag{20}$$

In terms of these new variables the model is

$$C_t + F_x = 0\,, \qquad F_t + 2\alpha F = -U^2 C_x\,. \tag{21}$$

The first equation is conservation of particles and the second equation is the flux-gradient relation.

Notice that in (21) Fick's law does not apply — the flux $F$ is not instantaneously related to the gradient $C_x$. Equation (21b), which might be called Cattaneo's law (see the 1989 review by Joseph and Preziosi), can be solved as a first-order differential equation for $F(x,t)$. Thus, the flux at $x$ is expressed as weighted integral over the past history of the gradient at $x$:

$$F(x,t) = -U^2 \int_{-\infty}^{t} e^{-2\alpha(t-t')} C_x(x,t')\,\mathrm{d}t'\,. \tag{22}$$

The flux has a "fading memory" of the gradient and the exponential in (22) is the fading factor which strongly weights the most recent values of the gradient.

Next, if we eliminate $F$ from (21), we obtain

$$C_{tt} + 2\alpha C_t - U^2 C_{xx} = 0\,. \tag{23}$$

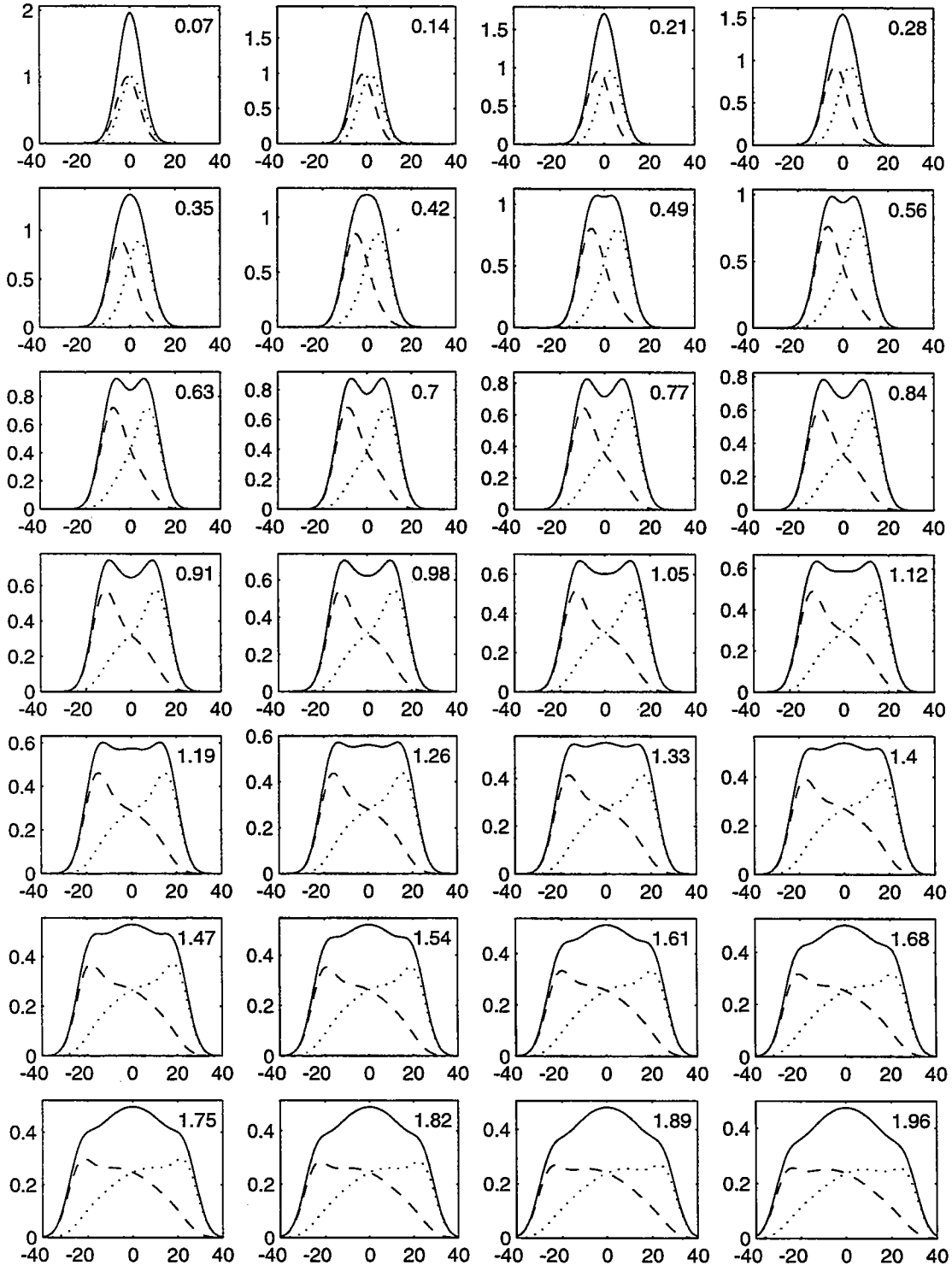


Figure 6: Solution of the telegraph equation. $\alpha t$ is in the top corner of the panel. At $t = 0$, $R = L = \exp(-x^2/50)$. The solid curve is $C = R + L$, and $R$ and $L$ are shown as dotted and dashed curves.

This is the telegraph equation; the diffusion equation is obtained only as an approximation which applies to the low frequency and wavenumber components of $C(x,t)$. On these large and slowly evolving scales one can neglect the term $C_{tt}$ in (23) and so obtain the approximation

$$C_t \approx DC_{xx}\,, \qquad D = \frac{U^2}{2\alpha}\,. \tag{24}$$

The diffusivity $D$ in (24) was anticipated in (13).

Figure 6 shows a numerical solution of (23) starting with an initial condition of the form

$$R(x,0) = L(x,0) = e^{-\mu^2 x^2}. \tag{25}$$

At small times the density $C$ develops a double peaked structure as the left and right going populations separate. This behaviour is transient, and at longer times the central part of the concentration relaxes to the well-known Gaussian solution of the diffusion equation.

According to (23) the disturbance travels at a finite speed: these are the "heat waves" discussed by Joseph and Preziosi (1989), and also evident in figure 6. The approximate diffusion equation (24) makes the unrealistic prediction that disturbances are propagated at infinite speed. This unphysical consequence of the diffusion equation motivated Cattaneo to propose (21b) as an alternative to Fick's law.

These considerations shows that one cannot blithely assert the validity of the diffusion equation (24) as an exact description of dispersion. The diffusion equation applies only as an approximate description of low frequencies and long wavelengths.

## 2.3 Discretization of the telegraph model

This section is a digression. Read on if you want to learn how to solve the telegraph equation using a simple numerical scheme. (This is how I drew figure 6.)

We reformulate the telegraph model in terms of discrete variables: divide the $x$-axis is divided into segments of length $\delta x$ separated by "scattering sites" at $x_n = n\delta x$. Time is also discretized in units of $\delta t$ so that $t = T\delta t$ where $T$ is an integer $T = 0, 1, 2\cdots$ The walkers move along the $x$–axis with a velocity that is either $+\delta x/\delta t$ or $-\delta x/\delta t$. In terms of the continuous model in (19)

$$U = \frac{\delta x}{\delta t}\,. \tag{26}$$

When a walker reaches the scattering site at $x_n = n\delta x$ he is "backwards scattered" or reflected with probability $b$ and "forward scattered" or transmitted with probability $1 - b$. Because the probability of a change in direction, $b$, is the same for left as for right moving walkers there is no mean velocity along the line.

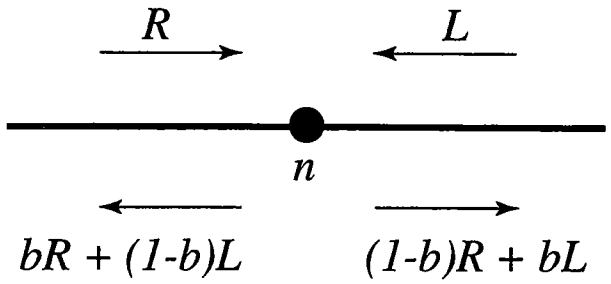


Figure 7: A steady state with constant flux, $f = U(1-b)(R-L)$, passing through site $n$.

Let $R_n(T)\delta x$ be the number of right walkers in the segment $n$, $n\delta x < x < (n+1)\delta x$. The number of left walkers, $L_n(T)\delta x$, is defined analogously. $R_n$ and $L_n$ are the discrete analogs of the continuous densities used in (19).

With these rules and definitions, the discrete evolution equations for the ensemble are

$$R_n(T) = (1-b)R_{n-1}(T-1) + bL_n(T-1)\,, \tag{27}$$

$$L_n(T) = (1-b)L_{n+1}(T-1) + bR_n(T-1)\,. \tag{28}$$

For instance, in the first equation above, the number of right movers in segment $n$ is equal to the number in segment $n-1$ at the previous time that successfully passed through scattering site $n$, plus the number of left movers previously in segment $n$ that were reflected at this same site. Figure 6 shows the result of iterating the discrete system above.

One exact solution of the difference equations above is

$$R_n = L_n = L_{n+1} = R_{n+1} = \cdots \tag{29}$$

This solution is steady: $R_n(T) = R_n(T-1)$. In fact, (29) is the discrete analog of the *equilibrium* solution of the diffusion equation. The distribution of walkers is spatially uniform with equal numbers going left and right in each interval and there are no concentration gradients. An individual walker is moving to and fro, but the ensemble is in steady state.

Next, we consider the *constant-flux* solution. In figure 7, $R$ right walkers impinge on site $n$ from the left and $L$ left walkers impinge on $n$ from the right. In steady state it must be that on the left of $n$ there are $bR + (1-b)L$ left walkers moving away, while to the right there are $bL + (1-b)R$ right walkers moving away. Thus the flux to the right of the site is

$$f = U\left[(1-b)\,R + bL\right] - UL = \frac{\delta x}{\delta t}(1-b)(R-L)\,, \tag{30}$$

where we have used $U = \delta x/\delta t$. Because there is a steady state, calculating the flux to the left of the site gives exactly the same result and so there is a nondivergent flux of walkers along the line.

Next, we can calculate the concentration difference across the site in figure 7. To the right of the site the total density of walkers is

$$c^+ = L + bL + (1-b)R, \tag{31}$$

while on the left the density is

$$c^- = R + bR + (1-b)L. \tag{32}$$

Combining (31) and (32) we have for the concentration jump across the site

$$\delta c = c^+ - c^- = 2b(L-R) \tag{33}$$

Thus, using (30), the flux-gradient relation in steady state is

$$f = -D\frac{\delta c}{\delta x}, \qquad D = \frac{(1-b)}{2b}\frac{(\delta x)^2}{\delta t}. \tag{34}$$

Does it seem obvious to you that the diffusivity should diverge as $b \to 0$? If you think of the diffusivity as the area under the correlation functions then this divergence should be intuitive. It is an instructive exercise to obtain $D$ in (34b) using Taylor's formula. (Hint: consider $N \gg 1$ right walkers which initially set out together. At $t = T\delta t$, after $T$ encounters with scattering sites, how many of these walkers have changed direction an even number of times, and how many odd?)

Comparing the equation above with our earlier expression for the diffusivity, $D$ in (13) and (24), we conclude that

$$\alpha\delta t = \frac{b}{1-b}. \tag{35}$$

Thus, with (26) and (35), we can express the parameters of the discrete model, $(\delta x, \delta t, b)$, in terms of the parameters characterizing the continuous model, $U$ and $\alpha$.

# 3 Age-stratified populations

The telegraph model from section 2 is *Markovian.* This means that each particle has a constant probability per unit time, $\alpha$, of switching direction. Thus, no matter how long a particle has been moving to the right (say), its probability of switching direction in the next $dt$ is always $\alpha dt$. Consequently an exponentially decreasing number of particles move coherently for long intervals and the telegraph model in (19) does not exhibit anomalous diffusion.

A satisfactory description of anomalous diffusion demands a non-Markovian model in which particles have some memory of their past motion. To obtain superdiffusion it is necessary that a right-moving particle is less and less likely to change direction as it spends more and more time moving right[1].

Such memory effects are implicit in the models discussed by Weeks et al (1996), and in several of the articles in the conference proceedings edited by Schlesinger, Zaslavsky & Frisch (1994). The stochastic models discussed in Schlesinger et al. draw heavily on statistical physics. In this lecture we are going to develop the theory from scratch using a formalism which is accessible to people whose background is in fluid mechanics. The climb begins with an excursion into the theory of age-stratified populations.

Consider a population of items with a finite lifetimes and a death rate which depends on age, $a$. For example, light bulbs in a large building, or the population of the United States. At time $t$ the age structure of the population is characterized by a density function for which $f(a,t)\mathrm{d}a$ is the number of items whose age is between $a$ and $a+\mathrm{d}a$. In terms of $f$, the total number of items in the population, $N(t)$, and the average age, $\bar{a}(t)$, are given by

$$N(t) = \int_0^\infty f(a,t)\,\mathrm{d}a, \qquad \bar{a}(t) \equiv N^{-1}\int_0^\infty a f(a,t)\,\mathrm{d}a. \tag{36}$$

The density function evolves according to

$$f_t + f_a + \alpha f = 0, \tag{37}$$

where $\alpha(a)$ is the death-rate. The term $f_a$ in (37) says that the population translates along the age-axis at a rate one year every year. To completely specify the problem we must supply an initial condition, and also a boundary condition at $a = 0$. The boundary condition at $a = 0$ has an obvious interpretation:

$$f(0,t) = \text{the birth (or replacement) rate.} \tag{38}$$

In the case of a population of people, the boundary condition above is a flux of babies into the system.

The Markovian limit is the special case in which $\alpha$ is independent of $a$. This model of $\alpha$ is unrealistic for both light-bulbs and people, though it might apply to a population of radioactive molecules. The Markovian case is very simple because one can integrate (37) over $a$ and obtain a closed equation for $N(t)$:

$$N_t + \alpha N = f(0,t)\,. \tag{39}$$

Thus if $\alpha$ is constant and we need only the total number of functional items at $t$ then we do not need to solve partial differential equations and deal with the age structure of the population.

[1]A popular metaphor for the Markovian case is radioactive decay: a molecule has a constant probability per unit time of decaying. As a metaphor for the non-Markovian case, imagine entering an enormous maze and then trying to find your way back to the entrance. The longer one has wandered, the less the chance of stumbling on the exit in the next d$t$.

## 3.1 The steady-state solution

As a first illustrative example, suppose that the replacement rate is adjusted so that $N$ is constant. (Janitors replace light bulbs as soon as they burn-out.) In this case the equilibrium solution of (37) is

$$f(a) = N\tau^{-1}\Psi(a), \tag{40}$$

where

$$\Psi(a) \equiv \exp\left(-\int_0^a \alpha(a')\,da'\right), \qquad \tau \equiv \int_0^\infty \Psi(a)\,da. \tag{41}$$

The function $\Psi$, and its integral $\tau$, will occur frequently in the sequel. Notice that the replacement rate is $f(0) = N/\tau$ and this suggests that $\tau$ should be the average lifetime of an item. On the other hand, $\tau$ will not usually be equal to $\bar{a}$ in (36). I suggest brooding on this "paradox" and, as an exercise, see if you can resolve this confusion to your satisfaction by the end of this section.

In (41) we assume that the death rate $\alpha(a)$ is such that as $a \to \infty$, $\Psi(a) \to 0$ fast enough to ensure that $\tau$ is finite. For instance, if $\alpha$ is constant (this is the Markovian case) then $\Psi(a) = \exp(-\alpha a)$ and $\tau = \bar{a} = 1/\alpha$.

If the death rate $\alpha$ decreases with age then the average liftetime $\tau$ might not be finite. For example, consider the specific model

$$\alpha = \frac{\nu}{\theta + a}, \qquad \Rightarrow \qquad \Psi(a) = \left(\frac{\theta}{\theta + a}\right)^\nu. \tag{42}$$

Provided that $\nu > 1$ then the integral of $\Psi(a)$ converges and $\tau = \theta/(\nu - 1)$.

If $\nu < 1$ then $\tau = \infty$ and there is no steady solution. To understand this curious result we must solve an initial value problem (see appendix A). Here we just remark that if $\nu < 1$ then the average lifetime of a bulb is infinite. Detailed solution of the initial value problem in appendix A shows that in this case the replacement rate is $f(0,t) \propto t^{\nu-1}$. That is, the total number of new bulbs which have been installed at time $t$ grows like $t^\nu \ll t$. The hypothetical manufacturer of lightbulbs with $\nu < 1$ is threatened with bankruptcy: sales decrease with time, even though every bulb eventually fails.

## 3.2 A cohort of babies

Imagine a cohort of babies leaving the maternity ward together, or a box of new lightbulbs shipped fresh from the factory. These items will function for varying amounts of time, and so we can speak of the PDF of lifetimes. We denote this PDF by $\mathcal{P}(a)$, and our goal is to relate $\mathcal{P}(a)$ to the death rate $\alpha(a)$.

Consider a group of $N$ items which all start with $a = 0$ at $t = 0$. What fraction of this cohort survives at $t > 0$? The surviving fraction is also the fraction of lifetimes longer than $t$ and so

$$\text{surviving fraction at } t \equiv \Psi(t) = \int_t^\infty \mathcal{P}(a)\,da. \tag{43}$$

To calculate the surviving fraction, we solve (37) with the initial and boundary conditions

$$f(a,0) = N\delta(a), \qquad f(0,t) = 0. \tag{44}$$

The solution of (37) and (44) is

$$f(a,t) = N\Psi(t)\delta(a-t)\,, \tag{45}$$

where $\Psi$ is defined in (41). Thus $\Psi(t)$ is the fraction of the cohort which is still alive at time $t$; we refer to $\Psi$ as the *survival function*.

It now follows from (43) that the PDF of lifetimes of new items is

$$\mathcal{P}(a) = -\Psi_a = \alpha\Psi. \tag{46}$$

The average lifetime, $\tau$, is given by the equivalent expressions:

$$\tau = \int_0^\infty a\mathcal{P}(a)\,da = -\int_0^\infty a\Psi_a\,da = \int_0^\infty \Psi(a)\,da. \tag{47}$$

Thus, as was suggested in the discussion following (41), to keep a population in equilibrium the replacement rate is equal to the size of the population, $N$, divided by the average lifetime of *new* items, $\tau$.

## 3.3 Extinction of a population

As a final example, suppose that at $t = 0$ we have the steady-state lightbulb population in (40). If the janitors then go on strike, so that bulbs burn out without replacement, then how many bulbs are still operating at $t > 0$? In this example we must solve (37) with the initial and boundary conditions that

$$f(a,0) = N\tau^{-1}\Psi(a), \qquad f(0,t) = 0\,. \tag{48}$$

The solution is

$$f(a,t) = H(a-t)N\tau^{-1}\Psi(a)\,, \tag{49}$$

where $H(a-t)$ is the step function. Thus the fraction of surviving bulbs at $t$ is

$$\Theta(t) = \tau^{-1} \int_t^\infty \Psi(a)\, da = \tau^{-1} \int_t^\infty (a-t)\mathcal{P}(a)\, da. \tag{50}$$

Using the specific model of $\alpha$ in (42), the surviving fraction is

$$\Theta(t) = (1+\theta^{-1}t)^{1-\nu}. \tag{51}$$

$\Theta(t)$ is the most slowly decaying function we have seen so far: as $t \to \infty$, $\Theta(t) \gg \Psi(t) \gg \mathcal{P}(t)$. This model may be relevant to the very slow extinction of professors once the supply of graduate students is cut-off.

Comparing the results in sections 3.1 and 3.2, we see that the steady state population in section 3.1 contains more long-lived items than are in a cohort of new items section 3.2. This means that the average lifetime of the light bulbs currently operating in the Empire State building is longer than the average lifetime of bulbs shipped from the factory. The reason is obvious: items with brief lifetimes fail quickly, and will likely be replaced with items whose lifetime is closer to the mean. Thus, fragile individuals are underrepresented in an operational population.

# 4 The generalized telegraph model

## 4.1 Formulation

Using the machinery from the previous section we now construct a generalization of the telegraph model which exhibits anomalous diffusion. In this generalization particles switch randomly between moving with $u(t) = +U$, $u(t) = 0$ and $u(t) = -U$. The transition probabilities between these states are functions of the time since the last transition. In other words, each particle carries an "age", $a$, which is the time elapsed since the particle transitioned into its present state. We denote the density of right moving particles at $(x,t)$, with age $a$, by $\mathcal{R}(a,x,t)$. For left-moving particles the density is $\mathcal{L}(a,t,x)$, and for the stationary particles the density is $\mathcal{S}(a,x,t)$. We refer to the left and right–movers collectively as "flying particles" while the stationary particles are "stickers".

The flying particles satisfy the conservational laws

$$\mathcal{R}_t + \mathcal{R}_a + U\mathcal{R}_x + \alpha_F \mathcal{R} = 0, \quad \mathcal{L}_t + \mathcal{L}_a - U\mathcal{L}_x + \alpha_F \mathcal{L} = 0\,, \tag{52}$$

while the sticking particles have

$$\mathcal{S}_t + \mathcal{S}_a + \alpha_S \mathcal{S} = 0\,. \tag{53}$$

The death rates of flying and sticking particles, $\alpha_F$ and $\alpha_S$ respectively, are functions of age $a$; it is through this device that particles have a memory of their previous

history. The price we pay for this nonMarkovian memory is that there are now three independent variables, $(a, t, x)$.

Stationary particles are born when left and right-moving particles die. And, conversely, when a stationary particle dies it is reborn as either a left moving particle or a right moving particle with equal probability. Notice that in order for a right-moving particle to become a left-moving particle it must pass through the intermediate state with $u = 0$. These karmic rules are enforced by boundary conditions at $a = 0$:

$$\mathcal{L}(0,t,x) = \mathcal{R}(0,t,x) = \frac{1}{2}\int_0^\infty \alpha_S(a)\mathcal{S}(a,t,x)\,\mathrm{d}a\,, \tag{54}$$

and

$$\mathcal{S}(0,t,x) = \int_0^\infty \alpha_F(a)\left[\mathcal{L}(a,t,x) + \mathcal{R}(a,t,x)\right]\,\mathrm{d}a\,. \tag{55}$$

Trajectories of particles moving with this generalized telegraph process are shown in figure 8.

The model we have formulated here is a generalization of the telegraph model in two ways. First, there are three states: left, right and stationary. This minor embellishment is motivated by the Texas experiments in which trapping in a vortex corresponds to the stationary particles. The nontrivial generalization is the introduction of the age variable used to capture memory effects. As an exercise, the student can show that if $\alpha_F$ and $\alpha_S$ are independent of $a$ then one can easily integrate over $a$ and reduce (52) through (55) to a three-state telegraph model. (This exercise shows how the boundary condition at $a = 0$ works.) As a sequel to this exercise, discuss $\alpha_S \to \infty$ and show that in this limit one obtains effectively a two-state telegraph model. Are you surprised that the diffusivity is given by (18)? That is, why do we recover model B, rather than model A, when the sojourns in the intermediate state $u = 0$ are very brief?

In order to model slowly fading velocity correlations and anomalous diffusion we use

$$\alpha_F(a) = \frac{\nu_F}{\theta_F + a}\,, \qquad \alpha_S(a) = \frac{\nu_S}{\theta_S + a}\,. \tag{56}$$

With the form above, the transition rates decrease as particles age. Numerical simulations of the three-state model using the transition rates in (56) show that many particles move in the same direction for a long time (see figure 9).

The main point of (56) is that if $a \gg 1$ then the transition rates $\alpha_F$ and $\alpha_S$ are proportional to $a^{-1}$. This inverse dependence on age ensures that the flying and sticking PDFs, $\mathcal{P}_F$ and $\mathcal{P}_S$ in (7), decay algebraically. Thus (56) incorporates important experimental information into the model[2]. One can make a dimensional argument in

[2]As far as scaling exponents are concerned, only the $a \gg 1$ structure of $\alpha_F$ and $\alpha_S$ matter. We use the specific functional form in (56) for simplicity.

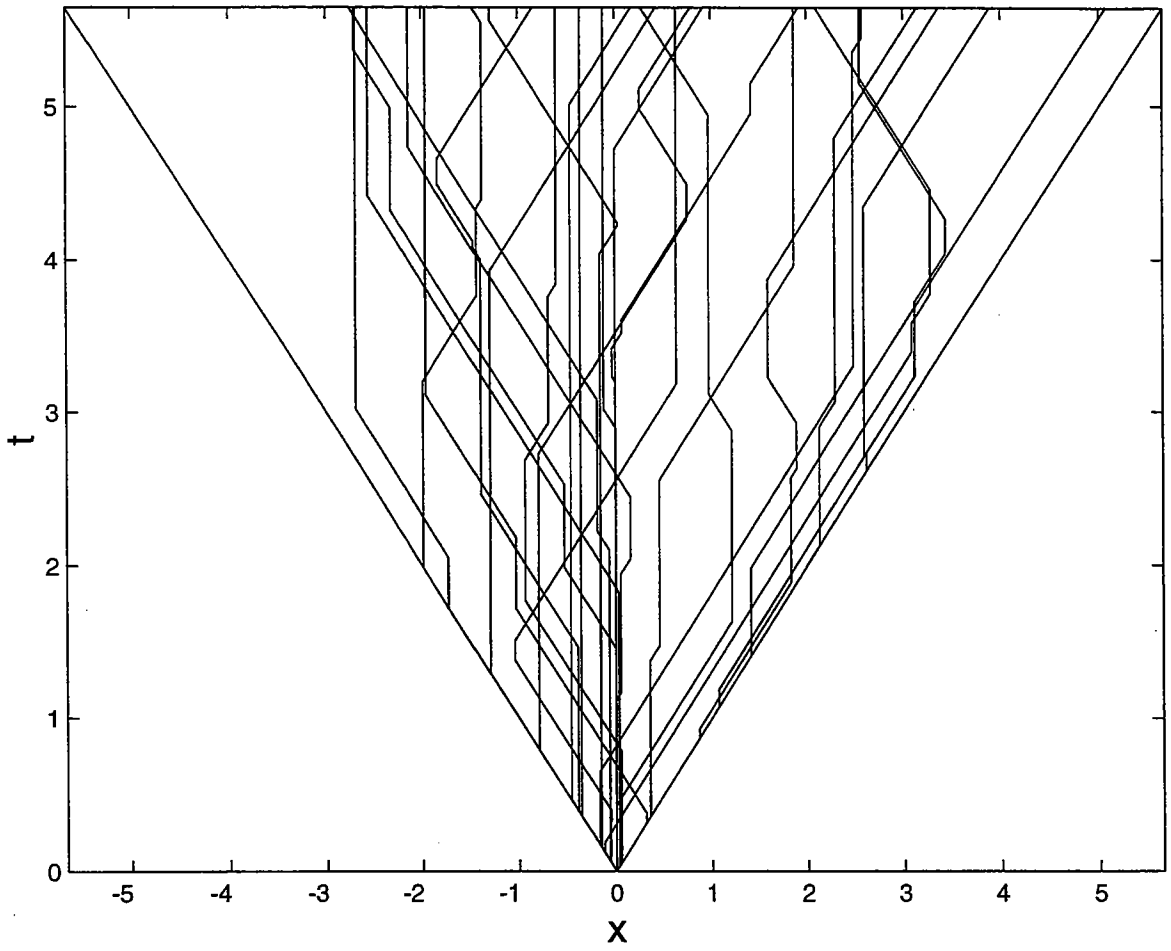


Figure 8: Trajectories of particles in the generalized telegraph random process. All particles are released from $x = 0$.

support of (56): $\alpha_F$ and $\alpha_S$ have the dimensions of inverse time. If the only time-scale relevant for long-lived particles is the particle age, $a$, then it follows that $\alpha_F$ and $\alpha_S$ are inversely proportional to $a$. We now show that the parameters $\nu_F$ and $\nu_S$ are easily related to the experimentally measured exponents $\mu_F$ and $\mu_S$ in (9).

## 4.2 The equilibrium solution

The system (52) through (55) has a solution which is homogeneous ($\partial_x = 0$) and steady ($\partial_t = 0$). This equilibrium solution is

$$\mathcal{R}(a,x,t) = \mathcal{L}(a,t,x) = r\Psi_F(a), \qquad \mathcal{S}(a,t,x) = 2r\Psi_S(a) \tag{57}$$

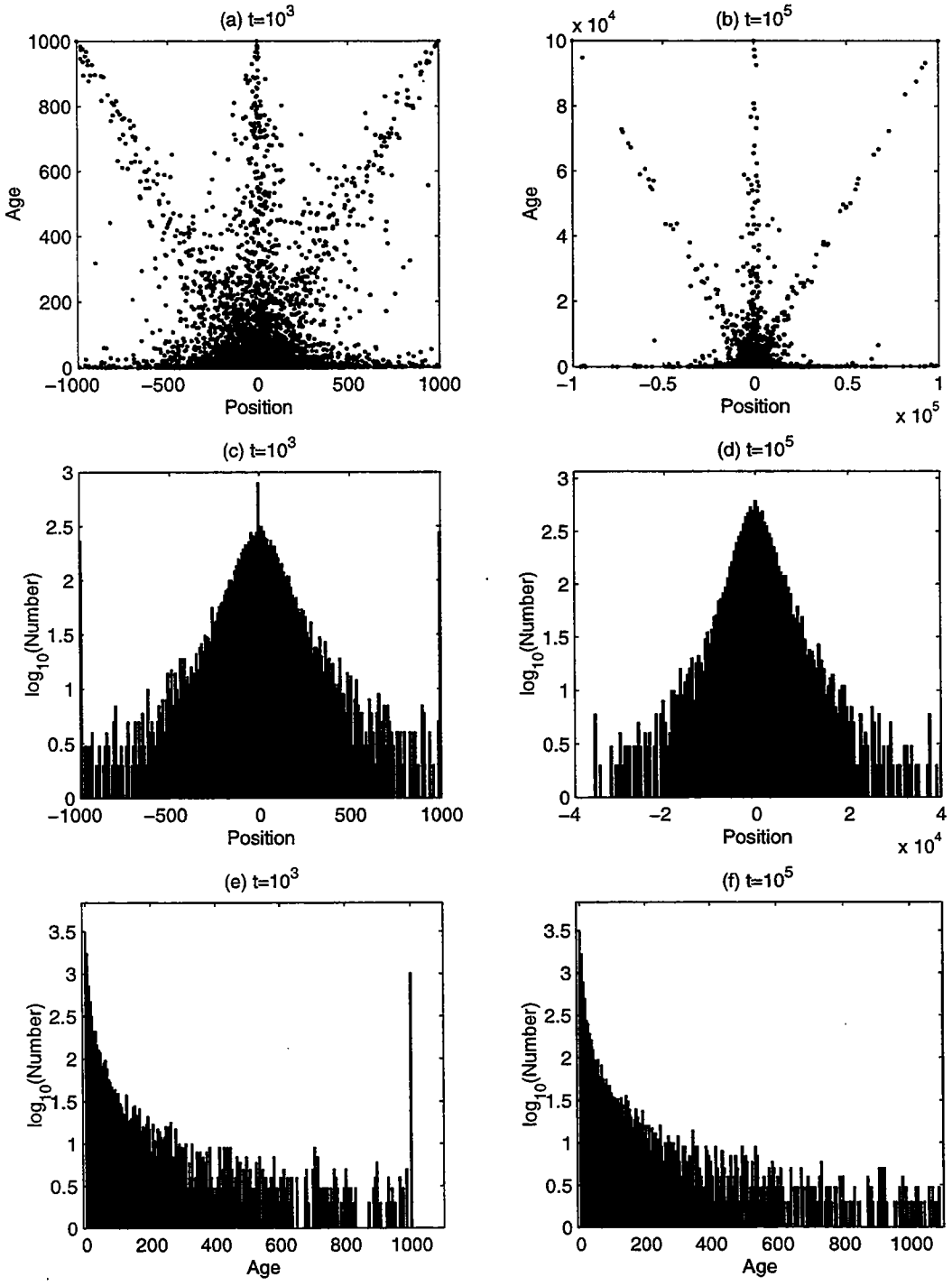


Figure 9: A simulation with $N = 10^4$ particles; $\alpha_F(a) = \alpha_S(a) \sim 1.35/a$. Upper panels: PDFs as a function of age and position show that there are many particles that either stick or move at a constant velocity for nearly the whole simualtion. Center panels: PDFs of the position of particles develop tails larger than Gaussians as time goes on. Lower panels: PDFs of the age of particles have a spike at large times, because there is a fraction of particles that never die.

where $\Psi_{F,S}(a)$ is

$$\Psi_{F,S}(a) \equiv \exp\left(-\int_0^a \alpha_{F,S}(a')\,\mathrm{d}a'\right). \tag{58}$$

The constant $r$ in (57) is the transition rate between the different states; $r$ is determined by the normalization condition:

$$N = 2r\,(\tau_S + \tau_F)\,, \qquad \tau_{F,S} \equiv \int_0^\infty \Psi_{F,S}(a)\,\mathrm{d}a\,. \tag{59}$$

We can use the results from section 3 to interpret $\tau_F$ and $\tau_S$ as the average lifetimes in the flying and sticking states respectively. Using (46), the PDF of lifetimes in those states is given by

$$\mathcal{P}_{F,S} = \alpha_{F,S}\Psi_{F,S}\,. \tag{60}$$

Using the expression in (56) for $\alpha_{F,S}$, we see that as $a \to \infty$, the survival functions decay algebraically with $\Psi_{F,S} \sim a^{-\nu_{F,S}}$, and so $\mathcal{P}_{F,S} \sim a^{-\nu_{F,S}-1}$. It follows that the exponents $\mu_F$ and $\mu_S$ defined in (9) are related to $\nu_F$ and $\nu_S$ by

$$\mu_{F,S} = \nu_{F,S} + 1\,. \tag{61}$$

We can summarize our arguments to this point by observing that the experiments provide the flying velocity, $U$, the average lifetime in the flying and sticking states, $\tau_{F,S}$, and the exponents $\mu_{F,S}$. These five experimental data determine the five parameters in the generalized telegraph model, namely $(U, \nu_{F,S}, \theta_{F,S})$.

## 4.3 Formulation of the initial value problem

Now that we have determined the model parameters using experimental constraints it is time to do some mathematics and use the model to predict the exponent $\gamma$ in (6). The simplest intial value problem we can consider is (52) through (55) with

$$[\mathcal{R}(a,0,x), \mathcal{S}(a,0,x), \mathcal{L}(a,0,x)] = r\,[\Psi_F(a), 2\Psi_S(a), \Psi_F(a)]\,\delta(x). \tag{62}$$

The constant $r$ is given in (57). Thus, the initial population has an equilibrium distribution of ages and is released at $x = 0$. Because of the symmetry between left and right moving particles

$$\mathcal{R}(a,t,x) = \mathcal{L}(a,t,-x), \qquad \mathcal{S}(a,t,x) = \mathcal{S}(a,t,-x). \tag{63}$$

Equation (63) greatly simplifies subsequent algebra.

One technical point (which I confess confuses me) is using the equilibrium age distribution as the initial condition in (62). This choice leads to simple calculations

below. And perhaps the gross details of the dispersion process, such the exponent $\gamma$, are independent of the initial distribution of ages? As an excercise I suggest solving the initial value problem using other initial conditions e.g., $\mathcal{R}(a,0,x) = \delta(a)\delta(t)$ etcetera. Are there any significant differences in the $t \to \infty$ structure of the solution?

Our strategy will be to obtain a closed hierarchy of spatial moments by multiplying the conservation laws (52) and (53) by $x^n$ and integrating over $x$. It is possible to solve the first few members of the hierarchy and show that if $\alpha_{F,S}$ and has the form in (56) with $1 < \nu < 2$ then as $t \to \infty$

$$\langle x^2 \rangle = \int_0^\infty \int_{-\infty}^\infty x^2 \left[\mathcal{R}(a,t,x) + \mathcal{S}(a,t,x) + \mathcal{L}(a,t,x)\right] \, dx da \propto t^{3-\nu_F} \, . \tag{64}$$

Before entering this calculation, we give a simple argument which suggests how the anomalous exponent $3 - \nu_F > 1$ arises in (64).

The variance $\langle x^2 \rangle$ in (64) can alternatively be written as

$$\langle x^2 \rangle = \frac{1}{N} \sum_{n=1}^{N} x_n^2. \tag{65}$$

At time $t > 0$ some of the $N$ particles will have moved coherently with unchanging velocity (either $+U$ or $-U$) ever since $t = 0$; half of these particles will be at $x = Ut$ and the other half at $x = -Ut$. These "coherent particles" each contribute a term $U^2t^2$ to the sum on the right hand side of (65). The number of coherent flying particles is just $\Theta(t)N$ where $\Theta(t)$ is given by (51) with $\nu$ replaced by $\nu_F$. Thus, because every term in the sum in (65) is positive, one has

$$\langle x^2 \rangle > \Theta(t) U^2 t^2 \sim U^2 \theta_F^{\nu-1} t^{3-\nu_F} \tag{66}$$

The inequality (66) has teeth only if $3 - \nu_F > 1$: then we learn that the coherent particles alone produce a superdiffusive contribution to the variance.

The argument above may suggest to you that superdiffusion is due solely to the few extreme particles which move without changes in direction. This is an overstatement: the lower bound in (66) is generously less than the exact result for $\langle x^2 \rangle$ which we obtain in the next section. Thus "nearly-coherent" particles, meaning particles which change direction only once or twice, also make a large contribution to the sum in (65). This is an essential point, because in their analysis of the experiments Solomon et al. discarded all coherent particles from the data set[3]. Thus the exponent measured by Solomon et al reflects only the contribution of nearly coherent particles.

[3]This drastic procedure is necessary because some fraction of the experime ntal particles are in integrable regions and will fly forever. Retaining all these particles will ultimately lead to the trival ballistic exponent $\gamma = 2$.

## 4.4 Solution of the initial value problem

This is a dry section which contains the details of the analytic calculation of $\langle x^2 \rangle$. The main point of interest here is that a lot of the algebra can be avoided by proving (75) below. (I suggest this as an exercise.)

The spatial moments are defined by

$$[\mathcal{R}_n(a,t), S_n(a,t), \mathcal{L}_n(a,t)] \equiv \int_{-\infty}^{\infty} x^n [\mathcal{R}(a,t,x), \mathcal{S}(a,t,x)\mathcal{L}(a,t,x)] \, dx, \tag{67}$$

Because of the symmetry in (63)

$$\mathcal{R}_n(a,t) = (-1)^n \mathcal{L}_n(a,t), \qquad \mathcal{S}_n(a,t) = 0 \text{ if } n \text{ is odd.} \tag{68}$$

The result above allows us to work exclusively with $\mathcal{R}_n$ and $\mathcal{S}_n$ while retaining full information about the distribution. Using the symmetry, the variance can be written as

$$\langle x^2 \rangle = \int_0^{\infty} 2R_2 + S_2 \, da \,. \tag{69}$$

The zeroth moment of (52) through (55), with the initial condition in (62) is

$$[\mathcal{R}_0(a,t), \mathcal{S}_0(a,t), \mathcal{L}_0(a,t)] = r\,[\Psi_F(a), 2\Psi_S(a), \Psi_F(a)] \,. \tag{70}$$

That is, the zeroth moment is just the equilibrium solution. (This is why using the equilibrium age distribution as the initial condition is so convenient.)

Using (68), the first spatial moment is $\mathcal{S}_1 = 0$, $\mathcal{L}_1(a,t) = -\mathcal{R}_1(a,t)$ and

$$\mathcal{R}_{1t} + \mathcal{R}_{1a} + \alpha_F \mathcal{R}_1 = Ur\Psi_F, \qquad \mathcal{R}_1(0,t) = 0, \qquad \mathcal{R}_1(a,0) = 0. \tag{71}$$

The solution of the initial value problem (71) is

$$\mathcal{R}_1(a,t) = Ur\Psi_F(a) \min(a,t). \tag{72}$$

The second moment equations are $\mathcal{L}_2 = \mathcal{R}_2$ and

$$\mathcal{R}_{2t} + \mathcal{R}_{2a} + \alpha_F \mathcal{R}_2 = 2U\mathcal{R}_1, \qquad \mathcal{S}_{2t} + \mathcal{S}_{2a} + \alpha_S \mathcal{S}_2 = 0, \tag{73}$$

with the $a = 0$ boundary condition that

$$2\mathcal{R}_2(0,t) = \int_0^{\infty} \alpha_S(a)\mathcal{S}_2(a,t) \, da, \qquad \mathcal{S}_2(0,t) = 2\int_0^{\infty} \alpha_F(a)\mathcal{R}_2(a,t) \, da. \tag{74}$$

To obtain the variance in (69) we do not need the complete solution of (73) and (74). Instead, after some judicious integration over $a$, one finds that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x^2 \rangle = 4U \int_0^{\infty} \mathcal{R}_1(a,t) \, da. \tag{75}$$

Substituting (72) into the result above we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x^2\rangle = 4U^2 r\left[\int_0^t a\Psi_F(a)\,\mathrm{d}a + t\int_t^\infty \Psi_F(a)\,\mathrm{d}a\right]. \tag{76}$$

If the right hand side of (76) approaches a constant as $t\to\infty$ then the variance grows diffusively. Otherwise there is anomalous diffusion.

With (76) in hand, one can easily determine if particular models of $\alpha_F$ and $\Psi_F$ lead to anomalous diffusion. For example, with the model in (56), evaluating the integrals in (42) gives a pleasant exact solution

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x^2\rangle = 4U^2 r\theta_F^2\left[\frac{(1+\tilde{t})^{2-\nu_F}}{(2-\nu_F)(\nu_F-1)} + \frac{1}{(\nu_F-1)(\nu_F-2)}\right], \tag{77}$$

where $\tilde{t}\equiv t/\theta$.

The asymptotic behaviour at large time depends crucially on $\nu_F$. If $\nu_F > 2$ then there is normal diffusion:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x^2\rangle \approx \frac{4U^2 r\theta_F^2}{(\nu_F-1)(\nu_F-2)} + O(t^{2-\nu_F}). \tag{78}$$

If $1<\nu_F<2$, there is superdiffusion

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x^2\rangle \approx \frac{4U^2 r\theta_F^2\tilde{t}^{2-\nu_F}}{(2-\nu_F)(\nu_F-1)} + O(1). \tag{79}$$

(At $\nu_F = 2$ there is a logarithmic term.)

Notice the minor role of $\alpha_S(a)$ in the solution above: if $\nu_S > 1$, so that the mean sticking time is finite, then the parameters $\nu_S$ and $\theta_S$ occur only in $r$.

## 4.5 An exercise for the diligent student

Consider an asymmetric two-state model

$$\mathcal{L}_t + \mathcal{L}_a + U_L\mathcal{L}_x + \alpha_L(a)\mathcal{L} = 0, \qquad \mathcal{R}_t + \mathcal{R}_x + U_R\mathcal{R}_x + \alpha_R(a)\mathcal{R} = 0, \tag{80}$$

with the boundary conditions

$$\mathcal{L}(0,t,x) = \int_0^\infty \alpha_R(a)\mathcal{R}(a,t,x)\,\mathrm{d}a, \qquad \mathcal{R}(0,t,x) = \int_0^\infty \alpha_L(a)\mathcal{L}(a,t,x)\,\mathrm{d}a. \tag{81}$$

Show that the average velocity is

$$\bar{U} = \frac{\tau_L U_L + \tau_R U_R}{\tau_L + \tau_R}, \qquad \tau_{L,R} \equiv \int_0^\infty \Psi_{L,R}(a)\,\mathrm{d}a, \tag{82}$$

where $\Psi_L$ and $\Psi_R$ are defined by analogy with (41). Show that the Laplace transform of the velocity autocorrelation function is given by

$$\hat{\text{corr}}(s) = U_{RMS}^2 \left[ \frac{1}{s} - \frac{\tau_L + \tau_R}{\tau_L \tau_R} \frac{(1-\hat{\psi}_L)(1-\hat{\psi}_R)}{s^2(1-\hat{\psi}_L\hat{\psi}_R)} \right], \tag{83}$$

where

$$U_{RMS}^2 \equiv \frac{\tau_L + \tau_R}{\sqrt{\tau_L \tau_R}} U_R U_L. \tag{84}$$

(If you use the moment method, you will need Laplace transforms to solve the integral equation which arises at $n = 1$.) Using the model

$$\alpha_{R,L}(a) = \frac{\nu_{R,L}}{\theta_{R,L} + a}, \tag{85}$$

perform an asymptotic analysis of (83) to identify the anomalous diffusion exponents which occur if either or both of $\nu_L$ and $\nu_R$ are less than 2.

# References


[1] D.D. Joseph and L. Preziosi. Heat waves. *Rev. Mod. Phys.*, 61:41–73, 1989.

[2] M. Schlesinger, G.M. Zaslavsky, and U. Frisch. *Lévy Flights and Related Topics in Physics.* Springer, Berlin, 1994.

[3] T.H. Solomon, E. Weeks, and H.L. Swinney. Chaotic advection in a two-dimensional flow: Lévy flights and anomlaous diffusion. *Physica D*, 76:70–84, 1994.

[4] G.I. Taylor. Diffusion by continuous movements. *Proc. London Math. Soc.*, 20:196–212, 1921.

[5] E.R. Weeks, J.S. Urbach, and H.L. Swinney. Anomalous diffusion in asymmetric random walks with a quasi-geostrophic example. *Physica D*, 97:291–310, 1996.


# A Solution of an initial value problem

In this appendix we discuss the issue raised at the end of section 3.1 and analyze a problem in which the death rate of old items is so small that the average lifetime $\tau$ is infinite. For example, this is the case $\nu < 1$ in (42). Specifically, consider the initial value problem posed by (37) with the initial and boundary conditions

$$f(a,0) = N\delta(a), \qquad f(0,t) = r(t). \tag{86}$$

In (86) the replacement rate $r(t)$ is determined by requiring that

$$N = \int_0^\infty f(a,t)\,\mathrm{d}a\,. \tag{87}$$

The solution of (37) and (86) is

$$f = N\Psi(t)\delta(a-t) + \Psi(a)r(t-a). \tag{88}$$

The first term on the right hand side of the equation above is the cohort of initial items aging and dying. The second term is influx of new items. Imposing (87) on (88), we obtain an integral equation for $r$:

$$N = N\Psi(t) + \int_0^t \Psi(a)r(t-a)\,\mathrm{d}a. \tag{89}$$

The integral relation above is known as the *renewal equation*

Because of the convolution in (89), the Laplace transform

$$[\hat{\Psi}(s), \hat{r}(s)] \equiv \int_0^\infty e^{-sa}[\Psi(a), r(a)]\,\mathrm{d}a\,, \tag{90}$$

is gratifying. In this way we find from (89) that

$$\hat{r} = N\frac{1-s\hat{\Psi}}{s\hat{\Psi}}. \tag{91}$$

The large-time behaviour of $r(t)$ can be obtained from (91) using standard asymptotic methods.

If $\alpha(a) \propto 1/a$ as $a \to \infty$, then the rightmost singularity of $\hat{\Psi}(s)$in the complex $s$-plane is the branch-point at $s = 0$. We show below in (94) through (97) that the structure of $\hat{\Psi}$ at this branch-point is

$$\Psi(s) = \varpi s^{\nu-1} + \tau + \cdots \tag{92}$$

If $\nu < 1$ then the singular term involving $s^{\nu-1}$ dominates the constant $\tau$ as $s \to 0$. In this case, from(91),

$$\hat{r}(s) \sim \frac{N}{\varpi s^\nu}, \qquad \Rightarrow \qquad r(t) \sim \frac{Nt^{\nu-1}}{\varpi\Gamma(\nu)}, \qquad \text{as} \qquad t \to \infty \tag{93}$$

Because $\nu < 1$ the replacement rate vanishes as $t \to \infty$.

To explain the small-$s$ expansion in (92), we use the model death-rate in (42), which produces the survival function

$$\Psi(a) = \left(\frac{\theta}{\theta+a}\right)^\nu. \tag{94}$$

The Laplace transform in (90) is then

$$\hat{\Psi}(s) = \theta^{\nu} s^{\nu-1} e^{\theta s} \Gamma(1-\nu, \theta s), \tag{95}$$

where $\Gamma(a, x)$ is the incomplete $\Gamma$-function defined by Abramowitz & Stegun in their article **6.5.3**. This Laplace transform can be rewritten as

$$\hat{\Psi}(s) = \theta^{\nu} s^{\nu-1} e^{\theta s} \Gamma(1-\nu) - \theta \Gamma(1-\nu) \sum_{n=0}^{\infty} \frac{(\theta s)^n}{\Gamma(2-\nu+n)}. \tag{96}$$

The form above is convenient because the singular terms containing $s^{\nu-1}$ are localized in the first function on the right hand side. When $s \ll 1$ the expansion of (96) is

$$\hat{\Psi}(s) = \theta^{\nu} \Gamma(1-\nu) s^{\nu-1} + \frac{\theta}{\nu - 1} + O(s, s^{\nu}), \tag{97}$$

which is the form assumed in (92).